(19) Canadian Intellectual Property Office
An Agency of Industry Canada

Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

(11) CA 2 454 127 (13) A1

(40) 23.01.2003
(43) 23.01.2003

(12)

(21) 2 454 127

(22) 05.07.2002

(51) Int. Cl.[7]: C12N 15/82, A01H 5/00

(85) 09.01.2004

(86) PCT/EP02/007527

(87) WO03/006660

(30) 101 33 407.9 US 13.07.2001
101 59 455.0 US 04.12.2001
102 07 582.4 US 22.02.2002

(71) SUNGENE GMBH & CO. KGAA,
Correnssstr. 3
06466, GATERSLEBEN, XX (DE).

(72) HILLEBRAND, HELKE (DE).
LEIN, WOLFGANG (DE).
HEIM, UTE (DE).
KUNZE, IRENE (DE).
GLICKMANN, ERIC (DE).
JOST, RICARDA (DE).
SONNEWALD, UWE (DE).
HELL, RUEDIGER (DE).
HERBERS, KARIN (DE).

(74) ROBIC

(54) CASSETTES D'EXPRESSION POUR L'EXPRESSION TRANSGENIQUE D'ACIDES NUCLEIQUES
(54) EXPRESSION CASSETTES FOR THE TRANSGENIC EXPRESSION OF NUCLEIC ACIDS

(57)

The invention relates to expression cassettes and vectors, which contain vegetable constitutive promoters and to the use of said expression cassettes or vectors for the transgenic expression of nucleic acid sequences preferably selection markers in organisms, preferably in plants. The invention also relates to transgenic plants that have been transformed using these expression cassettes or vectors, to cultures, parts or propagation products derived from said plants, and to the use of said plants for producing food and animal feed agents, seeds, pharmaceuticals, or fine chemicals.

Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2454127 A1 2003/01/23

(21) 2 454 127

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2002/07/05
(87) Date publication PCT/PCT Publication Date: 2003/01/23
(85) Entrée phase nationale/National Entry: 2004/01/09
(86) N° demande PCT/PCT Application No.: EP 2002/007527
(87) N° publication PCT/PCT Publication No.: 2003/006660
(30) Priorités/Priorities: 2001/07/13 (101 33 407.9) US;
2001/12/04 (101 59 455.0) US;
2002/02/22 (102 07 582.4) US

(51) Cl.Int.^{7}/Int.Cl.^{7} C12N 15/82, A01H 5/00

(71) Demandeur/Applicant:
SUNGENE GMBH & CO. KGAA, DE

(72) Inventeurs/Inventors:
HEIM, UTE, DE;
HILLEBRAND, HELKE, DE;
KUNZE, IRENE, DE;
HERBERS, KARIN, DE;
SONNEWALD, UWE, DE;
GLICKMANN, ERIC, DE;
LEIN, WOLFGANG, DE;
...

(74) Agent: ROBIC

(54) Titre : CASSETTES D'EXPRESSION POUR L'EXPRESSION TRANSGENIQUE D'ACIDES NUCLEIQUES
(54) Title: EXPRESSION CASSETTES FOR THE TRANSGENIC EXPRESSION OF NUCLEIC ACIDS

(57) Abrégé/Abstract:
The invention relates to expression cassettes and vectors, which contain vegetable constitutive promoters and to the use of said expression cassettes or vectors for the transgenic expression of nucleic acid sequences preferably selection markers in organisms, preferably in plants. The invention also relates to transgenic plants that have been transformed using these expression cassettes or vectors, to cultures, parts or propagation products derived from said plants, and to the use of said plants for producing food and animal feed agents, seeds, pharmaceuticals, or fine chemicals.

CA 2454127 A1 2003/01/23
(21) 2 454 127
(13) A1

(72) Inventeurs(suite)/Inventors(continued): HELL, RUEDIGER, DE; JOST, RICARDA, DE

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
23. Januar 2003 (23.01.2003)

PCT

(10) Internationale Veröffentlichungsnummer
WO 03/006660 A1

(51) Internationale Patentklassifikation[7]: C12N 15/82, A01H 5/00

(21) Internationales Aktenzeichen: PCT/EP02/07527

(22) Internationales Anmeldedatum:
5. Juli 2002 (05.07.2002)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:

101 33 407.9	13. Juli 2001 (13.07.2001)	US
101 59 455.0	4. Dezember 2001 (04.12.2001)	US
102 07 582.4	22. Februar 2002 (22.02.2002)	US

(71) Anmelder *(für alle Bestimmungsstaaten mit Ausnahme von US)*: **SUNGENE GMBH & CO. KGAA** [DE/DE]; Corrensstrasse 3, 06466 Gatersleben (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder *(nur für US)*: **HEIM, Ute** [DE/DE]; Wasserstrasse 1, 06466 Gatersleben (DE). **HILLEBRAND, Heike** [DE/DE]; B2, 11, 68159 Mannheim (DE). **KUNZE, Irene** [DE/DE]; Mühlenweg 11, 06466 Gatersleben (DE). **HERBERS, Karin** [DE/DE]; Am Hange 6, 06484 Quedlinburg (DE). **SONNEWALD, Uwe** [DE/DE]; Am Hange 6, 06484 Quedlinburg (DE). **GLICKMANN, Eric** [FR/DE]; Planstrasse 22, 06458 Hedersleben (DE). **LEIN, Wolfgang** [DE/DE]; Magdeburgstrasse 3, 38176 Cremlingen (DE). **HELL, Rüdiger** [DE/DE]; Brechtstrasse 8, 06484 Quedlinburg (DE). **JOST, Ricarda** [DE/DE]; Markt 12, 06484 Quedlinburg (DE).

(74) Anwalt: **WERNER, Frank**; Basf Aktiengesellschaft, 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten *(regional)*: ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:
— *mit internationalem Recherchenbericht*

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: EXPRESSION CASSETTES FOR THE TRANSGENIC EXPRESSION OF NUCLEIC ACIDS

(54) Bezeichnung: EXPRESSIONSKASSETTEN ZUR TRANSGENEN EXPRESSION VON NUKLEINSÄUREN

(57) Abstract: The invention relates to expression cassettes and vectors, which contain vegetable constitutive promoters and to the use of said expression cassettes or vectors for the transgenic expression of nucleic acid sequences preferably selection markers in organisms, preferably in plants. The invention also relates to transgenic plants that have been transformed using these expression cassettes or vectors, to cultures, parts or propagation products derived from said plants, and to the use of said plants for producing food and animal feed agents, seeds, pharmaceuticals, or fine chemicals.

(57) Zusammenfassung: Die Erfindung betrifft Expressionskassetten und Vektoren, die pflanzliche konstitutive Promotoren enthalten, sowie die Verwendung dieser Expressionskassetten oder Vektoren zur transgenen Expression von Nukleinsäuresequenzen bevorzugt Selektionsmarkern in Organismen, bevorzugt in Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder Vermehrungsgut, sowie die Verwendung derselben zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien.

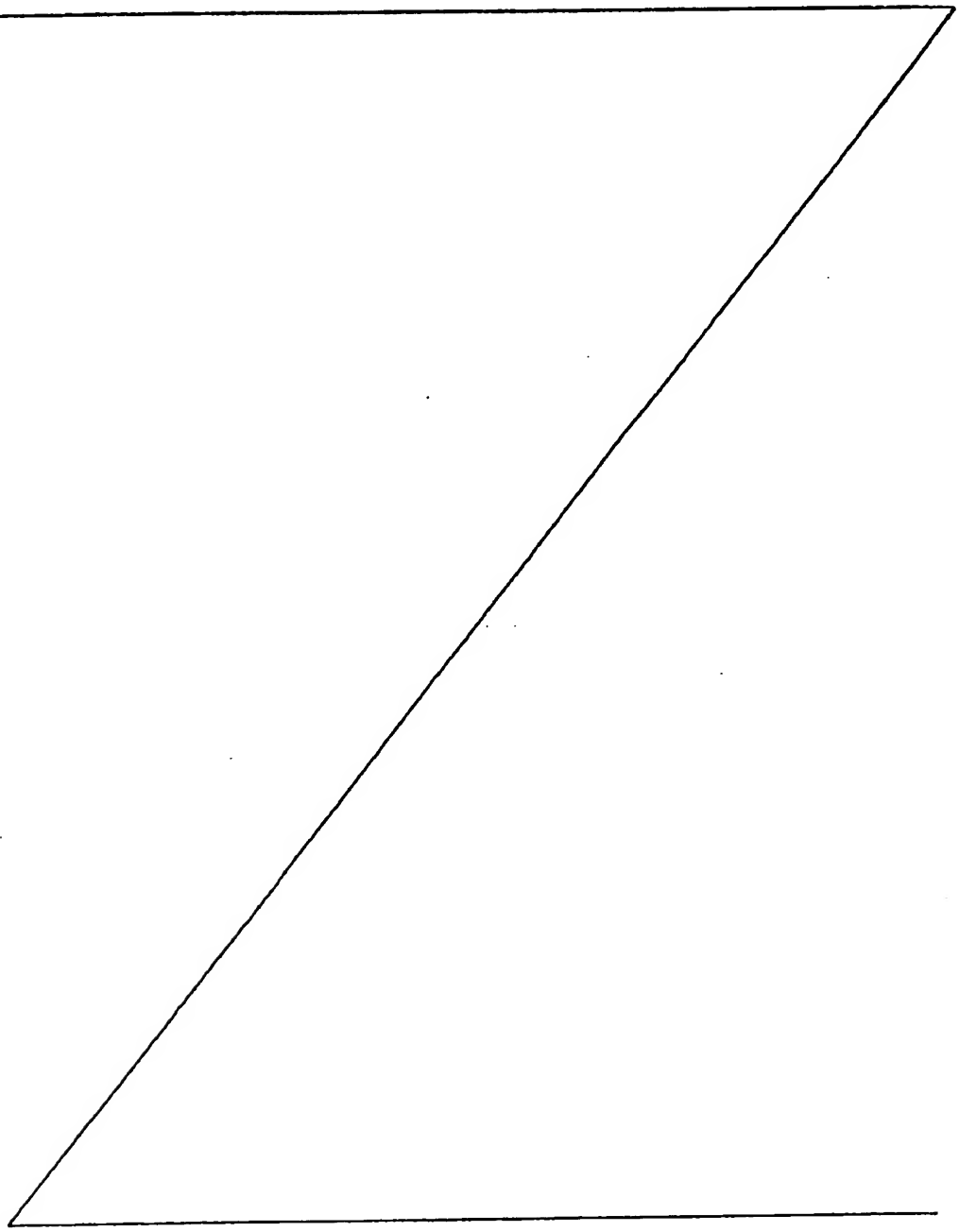
WO 03/006660 A1

EXPRESSION CASSETTES FOR THE TRANSGENIC EXPRESSION OF NUCLEIC ACIDS

The invention relates to expression cassettes and vectors which contain constitutive promoters of plants and to the use of said expression cassettes or vectors for transgenic expression of nucleic acid sequences, preferably selection markers, in organisms, preferably in plants. The invention further relates to transgenic plants which have been transformed with said expression cassettes or vectors, to cultures, parts or propagation material derived therefrom and also to the use of same for the production of food- and feedstuffs, seed, pharmaceuticals or fine chemicals.

The aim of biotechnological studies on plants is the preparation of plants having improved properties, for example to increase agricultural productivity. The preparation of transgenic plants is a fundamental technique in plant biotechnology and thus an indispensible prerequisite for basic research on plants in order for the preparation of plants having improved novel properties for agriculture, for improving the quality of foodstuffs or for the production of particular chemicals or pharmaceuticals (Dunwell JM, J Exp Bot. 2000;51 Spec No:487-96). The natural defence mechanisms of the plant, for example against pathogens, are often inadequate. The introduction of foreign genes from plants, animals or microbial sources can enhance the defence. Examples are the protection against insects feeding on tobacco by expression of the *Bacillus thuringiensis* endotoxin under the control of the 35 S CaMV promoter (Vaeck et al. (1987) Nature 328:33-37) and the protection of tobacco against fungal infection by expression of a chitinase from beans under the control of the CaMV promoter (Broglie et al. (1991) Science 254:1194-1197). It is furthermore possible to achieve resistance to herbicides by introducing foreign genes, thereby optimizing the cultivation conditions and reducing crop losses (Ott KH et al. (1996) J Mol Biol 263(2):359-368). The quality of the products may also be improved. Thus it is possible, for example, to increase the shelf life and storability of crop products by inactivating particular maturation genes. This was demonstrated, for example, by inactivating polygalacturonase in tomatoes (Hamilton AJ et al.(1995) Curr Top Microbiol Immunol 197:77-89).

A basic prerequisite for transgenic expression of particular genes in plants is the provision of plant-specific promoters. Various plant promoters are known. It is possible to distinguish between constitutive promoters which enable expression in various parts of a plant, which is only slightly restricted in terms of



location and time, and specific promoters which allow expression only in particular parts or cells of a plant (e.g. root, seeds, pollen, leaves, etc.) or only at particular times during development. Constitutive promoters are used, for example, for expressing "selection markers". Selection markers (e.g. antibiotic or herbicidal resistance genes) permit filtering the transformation event out of the multiplicity of untransformed but otherwise identical individual plants.

Constitutive promoters active in plants have beenm written [sic] relatively rarely up to now. Promoters to be mentioned are the *Agrobacterium tumefaciens*, TR double promoter, the promoters of the vacuolar ATPase subunits or the promoter of a proline-rich wheat protein (WO 91/13991) and also the Ppc1 promoter *Mesembryanthemum cryctallinum* (Cushman et al. (1993) Plant Mol Biol 21:561-566).

The constitutive promoters which are currently the predominantly used promoters in plants are almost exclusively viral promoters or promoters isolated from *Agrobacterium*. In detail, these are the nopaline synthase (*nos*) promoter (Shaw et al. (1984) Nucleic Acids Res. 12(20):7831-7846), the mannopine synthase (*mas*) promoter (Comai et al. (1990) Plant Mol Biol 15 (3):373-381) and the octopine synthase (*ocs*) promoter (Leisner and Gelvin (1988) Proc Natl Acad Sci USA 85(5):2553-2557) from *Agrobacterium tumefaciens* and the CaMV35S promoter from cauliflower mosaic virus. The latter is the most frequently used promoter in expression systems with ubiquitous and continuous expression (Odell et al.(1985) Nature 313:810-812; Battraw and Hall (1990) Plant Mol Biol 15:527-538; Benfey et al. (1990) EMBO J 9(69):1677-1684; US 5,612,472). However, the CaMV 35S promoter which is frequently applied as constitutive promoter exhibits variations in its activity in different plants and in different tissues of the same plant (Atanassova et al. (1998) Plant Mol Biol 37:275-85; Battraw and Hall (1990) Plant Mol Biol 15:527-538; Holtorf et al. (1995) Plant Mol Biol 29:637-646; Jefferson et al. (1987) EMBO J 6:3901-3907). A further disadvantage of the 35S promoter is a change in transgene expression in the case of an infection with cauliflower mosaic virus and its typical pathogenic variants. Thus, plants expressing the BAR gene under the control of the 35S promoter are no longer resistant after infection with the virus which typically occurs in nature (Al-Kaff et al. (2000) Natur Biotechnology 18:995-99).

From the range of viral promoters, the sugarcane bacilliform badnavirus (ScBV) which imparts an expression pattern similar to that of CamV has been described as an alternative to the CaMV 35S promoter (Schenk et al. (1999) Plant Mol Biol 39(6):1221-1230). The activity of the ScBV promoter was analyzed in transient expression analyses using various dicotyledonous plants, including *Nicotiana tabacum* and *N. benthamiana*, sunflower and oilseed rape, and monocotyledonous plants, here in the form of banana, corn and millet. In the transient analyses in corn, the ScBV promoter-mediated expression level was comparable to that of the ubiquitin promoter from corn (see below). Furthermore, the ScBV promoter-mediated rate of expression was assayed in transgenic banana and tobacco plants and displayed in both plant species essentially constitutive expression.

Common promoters for expressing selection markers in plants are especially the *nos* promoter, or else the *mas* promoter and *ocs* promoter, all of which have been isolated from *Agrobacterium* strains.

The use of viral sequences is often met with great reservations on the part of the consumer. These doubts are fed not least by studies which question the safety of the 35S CaMV promoter, owing to a possible horizontal gene transfer due to a recombination hot spot (Ho MW et al. (1999) Microbial Ecology in Health and Disease 11:194-197; Cummins J et al. (2000) Nature Biotechnology 18:363). It is therefore an aim of future biotechnological studies on plants to replace viral genetic elements by plant regulatory elements in order to keep as closely as possible to the plant system.

Owing to the prevailing doubts with regard to viral promoters, there are extensive efforts to replace said promoters by plant promoters. However, a promoter of plant origin, which is comparable to the viral elements, has not been described as yet.

What has been described, is a plant ubiquitin promoter from *Arabidopsis thaliana* (Callis et al.(1990) J Biol Chem 265:12486-12493; Holtorf S et al. (1995) Plant Mol Biol 29:637-747). Contrary to the findings in the articles mentioned, some studies revealed that the *Arabidopsis* ubiquitin promoter is unsuitable for expressing selection marker genes and that, for this reason, its general applicability must be called into question (see comparative examples 1 and 3).

The expression pattern mediated by the corn ubiquitin promoter has been described for the Ubi-1 and Ubi-2 promoters from corn (Christensen et al. (1992) Plant Mol Biol 18(4):675-689). While the Ubi-1 promoter has good expression activity in corn and other monocotyledonous plants, it exhibits in dicotyledonous tobacco plants only 10% of the activity which had been achieved in comparable experiments using the viral 35S promoter. It was furthermore shown that the corn Ubi-1 promoter is suitable for overexpression of genes in monocotyledonous plant systems and, in addition, is sufficiently strong in order to mediate a herbicidal resistance via the expression of selection markers (Christensen and Quail (1996) Transgenic Res 5(3):213-218). The Ubi-1 promoter proved unsuitable for dicotyledonous expression systems.

A comparison of the organ specificity and strength of various constitutive promoters was carried out by Holtorf (Holtorf et al. (1995) Plant Mol Biol 29(4):637-646) on the basis of stably transformed *Arabidopsis* plants. The study comprised, interalia, the CaMV35S promoter, the leaf-specific thionine promoter from barley and the *Arabidopsis* ubiquitin promoter (UBQ1). The CaMV35S promoter exhibited the highest rate of expression. On the basis of using an additional translational enhancer (TMV omega element), it was possible to increase the rate of expression of the promoter by a factor of two to three with unchanged organ specificity. The leaf-specific thionine promoter from barley was inactive in the majority of transformed lines, while the UBQ1 promoter from *Arabidopsis* resulted in medium rates of expression.

McElroy and colleagues reported a construct for transforming monocotyledonous plants, which is based on the rice actin 1 (Act1) promoter (McElroy et al. (1991) Mol Gen Genet 231:150-160). Overall, it was concluded from the afore-described studies that the Act1 promoter-based expression vectors are suitable for controlling a sufficiently strong and constitutive expression of foreign DNA in transformed cells of monocotyledonous plants.

Another constitutive promoter which has been described is the promoter of an S-adenosyl-L-methionine synthetase (WO 00/37662). A disadvantage here is especially a dependence of the strength of expression on the methionine concentration (see WO 00/37662; Fig. 7).

WO 99/31258 describes chimeric constitutive plant promoters which are composed of various elements of various promoters with complementary expression patterns so that combination of

individual tissue specificities additively results in a constitutive expression pattern.

Ferredoxin NADPH oxidoreductase (FNR) is a protein of the electron transport chain and reduces NADP+ to NADPH. Experiments in spinach using the spinach FNR promoter fused to the GUS gene hint at a light-inducible element in the FNR promoter (Oelmüller et al. (1993) Mol. Gen. Genet. 237:261-72). Owing to its function, a strictly leaf-specific expression pattern would have been expected for the promoter. Owing to the tissue-dependent expression pattern, the promoter would be poorly suited to expressing selection markers. Here, a selection in all tissue parts, if possible, is required in order to ensure efficient selection.

Owing to its function during photosynthesis, the promoter of the triose phosphate translocator (TPT) should be mainly leaf-specific. The cDNAs from potato (Schulz et al. (1993) Mol Gen Genet 238:357-61), cauliflower (Fischer et al. (1997) Plant Cell 9:453-62), oilseed rape (WO 97/25346) and corn Kammerer B (1998) The Plant Cell 10:105-117) have been described. Kammerer et al. demonstrate that TPT mRNA expression in corn is strong in the leaves and the stamen. In contrast, no expression was observed in the stem or in the roots. Owing to the tissue-dependent expression pattern, the promoter would be poorly suited to expressing selection markers. Here, a selection in all tissue parts, if possible, is required in order to ensure efficient selection.

The "constitutive" promoters described in the prior art have one or more of the following disadvantages:

1. Inadequate homogeneity of expression:
   The known "constitutive" promoters frequently display a different level of expression, depending on the type of tissue or cell. Moreover, the expression property is often highly dependent on the site of insertion into the host genome. As a consequence of this, the effects to be obtained by heterologous expression cannot be achieved to the same extent homogeneously in the plant. Under or over dosages may occur. This may have an adverse effect on plant growth or plant value.

2. Inadequate time profile:
The "constitutive" promoters known in the prior art often exhibit a nonconsistent activity during the development of a tissue. As a result, it is not possible, for example, to achieve desirable effects (such as selection) in the early phase of somatic embryogenesis which would be advantageous, especially here, due to the sensitivity of the embryo to in vitro conditions and stress factors.

3. Inadequate applicability to many plant species:
The "constitutive" promoters described in the prior art are often not active in the same way in all species.

4. If a plurality of expression cassettes with in each case the same "constitutive" promoter are present in an organism, interactions between said expression cassettes and even switching-off (gene silencing) of individual expression cassettes may occur (Mette et al. (1999) EMBO J. 18:241-248).

5. Promoters of viral origin may be influenced by virus infections of the transgenic plant and may then no longer express the desired property (Al-Kaff et al. (2000) Natur Biotechnology 18:995-99).

6. The public acceptance toward the use of promoters and elements from plant systems is higher than for viral systems.

7. The number of promoters suitable for expressing selection markers in plants is low and said promoters are usually of viral or bacterial origin.

8. Pollen/anther expression: The promoters mentioned (such as, for example, 35S CaMV) exhibit strong activity in the pollen or in the anthers. This may have disadvantageous effects on the environment. Thus, unspecific expression of *Bacillus thuringiensis* endotoxins resulted not only in the desired effect on feeding insects due to expression in the root but also, due to expression in the pollen, in considerable damage in the population of the Monarch butterfly which feeds predominantly on the pollen (Losey JE et al. (1999) Nature 399, 214).

An ideal constitutive promoter should have as many of the following properties as possible:

a) a gene expression which is as homogeneous as possible with regard to location and time, i.e. an expression in as many cell types or tissues of an organism as possible during the various phases of the developmental cycle. Furthermore, an efficient selection in differentiated cells (various callus phases) from a tissue culture and other developmental stages suitable for tissue culture is desired.

b) An applicability to various plant species, which is as broad as possible, and applicability to species in which it is not possible to achieve any expression using the "constitutive" promoters known to date.

c) In order to combine a plurality of transgenes in one plant, it is desirable to carry out a plurality of transformations in succession or to use constructs with a plurality of promoter cassettes, but without generating silencing effects due to the multiple use of identical regulatory sequences.

d) A plant origin in order to avoid problems of acceptance by the consumer and possible problems of future approval.

e) Secondary activities of a promoter in the anthers/pollen are undesirable, for example in order to avoid environmental damage (see above).

It is an object of the present invention to provide regulatory sequences of plants, which fulfill as many of the abovementioned properties as possible and which mediate especially a ubiquitous and development-independent (constitutive) expression of a nucleic acid sequence to be expressed which preferably codes for a selection marker. Despite various plant promoters for which a constitutive expression at least in individual species is claimed, no promoter having the desired properties listed above has been described up to now. It was therefore the object to identify appropriate promoters.

We have found that this object is achieved by providing expression cassettes based on the promoters of a putative ferredoxin gene (pFD "putative ferredoxin" hereinbelow) from *Arabidopsis thaliana*, of the ferredoxin NADPH oxidoreductase (FNR hereinbelow) gene from *Arabidopsis thaliana* and of the triose phosphate translocator (TPT) gene from *Arabidopsis thaliana*:

1.) Promotor of a putative ferredoxin (pFD) from *Arabidopsis thaliana*

During analysis of the *Arabidopsis genome*, the ORF of a putative ferredoxin gene was identified. The isolated 836 bp 5'- flanking sequence fused to the *Glucuronidase* gene surprisingly exhibited a constitutive expression pattern in transgenic tobacco. The sequence corresponds to a sequence section on *Arabidopsis thaliana* chromosome 4, as it has been deposited at GenBank under Acc. No. Z97337 (Version Z97337.2; base pair 85117 to 85952; the gene starting from bp 85953 is annotated "strong similarity to ferredoxin [2Fe-2S] I, Nostoc muscorum"). (The gene is not to be confused with the *A. thaliana* gene for preferredoxin annotated under GenBank Acc.-NoAcc. No: X51370; Vorst O et al. (1990) Plant Mol Biol 14(4):491-499).

Only a weak activity was detected in the anthers/pollen of the closed flower buds and no activity whatsoever was detected in mature flowers. Contrary to the reservations, derived from the findings in the literature, toward a suitability of the promoter for effective expression of selection markers (for example, owing to the suspected leaf specificity or function in the photosynthetic electron transport), it was possible to demonstrate a highly efficient selection by combination with, for example, the homolog resistance gene (nptII).

2.) Ferredoxin NADPH oxidoreductase (FNR) promoter from *Arabidopsis thaliana*

Starting from the information on FNR-encoding cDNA from *N. tabacum* (GenBank Acc. No.: Y14032) the Arabidopis data base was screened for a homologous gene. Primers were synthesized according to said sequence information. The promoter amplified via PCR from *Arabidopsis thaliana* genomic DNA (635 bp), of which a leaf-specific expression was expected, exhibited in transgenic tobacco plants a surprisingly ubiquitous and insertion site-independent expression.

The promoter sequence partly corresponds to a sequence section on *Arabidopsis thaliana* chromosome 5, as it is deposited at GenBank under Acc. No. AB011474 (Version AB011474.1 from 12.27.2000; base pair 70127 to 69493; the gene starting at bp 69492 is annotated with "ferredoxin-NADP+ reductase").

No activity was detected in the pollen. Contrary to the reservations, derived from the findings in the literature, toward a suitability of the promoter for effective expression of selection markers (for example, owing to the suspected leaf specificity or function in the photosynthetic electron transport), it was possible to demonstrate a highly efficient selection by combination with, for example, the phosphinothricin resistance gene (bar/pat).

The nondetectable activity of the FNR promoter in seeds allows a use for the expression of genes whose gene products are desired in other parts of the plant and are unwanted in the seeds. For example, pests can be repelled by expressing appropriate toxins such as, for example, *Bacillus thuringiensis* crystal proteins. Thus it is possible to achieve in potatoes expression in the plant organs above the ground (and thus, for example, a repulsion of pests such as the potato beetle) without simultaneous expression in the tuber which is used as food or animal feed, and this could increase the suitability and acceptance.

3.) Triose phosphate translocator (TPT) promoter from *Arabidopsis thaliana*

A 2038bp PCR fragment was amplified, starting from *Arabidopsis* GenBank data of chromosome V, clone MCL19. The promoter sequence partly corresponds to a sequence section on *Arabidopsis thaliana* chromosome 5, as it is deposited with GenBank under Acc. No. AB006698 (Version AB006698.1 from 12.27.2000; base pair 53242 to 55281; the gene starting at bp 55282 is annotated with "phosphate/triose-phosphate translocator").

Surprisingly, transgenic tobacco plants exhibited not only a high activity in numerous parts of the plant. No activity was detected in the pollen. Contrary to the reservations, derived from the findings in the literature, toward a suitability of the promoter for effective expression of selection markers (for example, owing to the suspected leaf specificity), it was possible to demonstrate a highly efficient selection by combination with, for example, the phosphinothricin resistance gene (bar/pat).

The ubiquitous expression pattern, but especially also the ability of the TPT promoter regarding the expression of selection markers, comes as a great surprise for the skilled worker, since the triosephosphate translocator is responsible

for the exchange of C3 sugar phosphates between the cytosol and the plastids in photosynthetic leaves. The TPT is located in the inner chloroplast membrane. Colorless plastids typically contain a hexose transporter via which C6-sugar phosphates are exchanged. It is not to be expected that such genes are active in the early callus and embryogenesis stages (Stitt (1997) Plant Metabolism, 2nd ed., Dennis eds. Longman Press, Harlow, UK, 382-400).

The pFD, FNR and TPT promoters proved to be sufficiently strong in order to express nucleic acid sequences, in particular selection marker genes, successfully. Furthermore, various deletion variants of the abovementioned promoters, in particular a truncated variant of the pFD promoter (699 bp) and of the TPT promoter (1318 bp), proved suitable for ensuring the expression of, for example, selection markers such as the homolog resistance (nptII).

Furthermore, the *Arabidopsis thaliana* ubiquitin promoter (Holtorf et al. (1995) Plant Mol Biol 29:637-646) and the squalene synthase promoter (Kribii et al. (1997) Eur J Biochem 249:61-69) were studied within the framework of the studies mentioned, both of which, however, were surprisingly unsuitable for mediating selection marker gene expression although the literature data of the ubiquitin promoters from monocotyledons (see above) had led to the assumption that in particular the ubiquitin promoter of a dicotyledonous plant should have worked as a promoter of a selection marker system (see comparative examples 1 and 3). A similar statement applies to the squalene synthase promoter whose characterization had led to the expectation that it would be possible to achieve sufficiently high rates of expression for the successful control of a selection marker gene (Del Arco and Boronat (1999) 4th European Symposium on Plant Isoprenoids, 21.-4.23.1999, Barcelona, Spain) (see comparative examples 2 and 3).

The present invention therefore relates firstly to expression cassettes for transgenic expression of nucleic acids, comprising

a) a promoter according to SEQ ID No: 1, 2 or 3,

b) a functional equivalent or equivalent fragment of a), which essentially possesses the same promoter activity as a),

a) or b) being functionally linked to a nucleic acid sequence to be expressed transgenically.

The invention further relates to methods for transgenic expression of nucleic acids, wherein a nucleic acid sequence which is functionally linked to

a) a promoter according to SEQ ID NO: 1, 2 or 3 or

b) a functional equivalent or equivalent fragment of a) which essentially possesses the same promoter activities as a),

is expressed transgenically.

Expression comprises transcription of the nucleic acid sequence to be expressed transgenically but may also include, in the case of an open reading frame in sense orientation, translation of the transcribed RNA of the nucleic acid sequence to be expressed transgenically into a corresponding polypeptide.

An expression cassette for transgenic expression of nucleic acids or a method for transgenic expression of nucleic acids comprises all those constructions produced by genetic methods or methods in which either

a) a promoter according to SEQ ID No: 1, 2 or 3 or a functional equivalent or equivalent fragment thereof, or

b) the nucleic acid sequence to be expressed, or

c) (a) and (b)

are not present in their natural genetic environment (i.e. at their natural chromosomal locus) or have been modified by genetic methods, and said modification may be, by way of example, a substitution, addition, deletion, inversion or insertion of one or more nucleotide residues.

The expression cassettes of the invention, vectors derived from them or the methods of the invention may comprise functional equivalents of the promoter sequences described under SEQ ID No: 1, 2 or 3. Functionally equivalent sequences also comprise all those sequences which are derived from the complementary counter [sic] strand of the sequences defined by SEQ ID NO: 1, 2 or 3 and which have essentially the same promoter activity.

Functional equivalents with respect to the promoters of the invention means in particular natural or artificial mutations of the promoter sequences described under SEQ ID No: 1, 2 or 3 and

of the homologs thereof from other plant genera and species, which furthermore have essentially the same promoter activity.

A promoter activity is essentially referred to as identical, if the transcription of a particular gene to be expressed under the control of a particular promoter derived from SEQ ID NO: 1, 2 or 3 under otherwise unchanged conditions has a location within the plant, which is at least 50%, preferably at least 70%, particularly preferably at least 90%, very particularly preferably at least 95%, congruent with that of a comparative expression obtained using a promoter described by SEQ ID NO: 1, 2 or 3. In this connection, the expression level may deviate both downward and upward in comparison to a comparative value. In this connection, preference is given to those sequences whose expression level, measured on the basis of the transcribed mRNA or the subsequently translated protein, differs quantitatively by not more than 50%, preferably 25%, particularly preferably 10%, from a comparative value obtained using a promoter described by SEQ ID NO: 1, 2 or 3, under otherwise unchanged conditions. Particular preference is given to those sequences whose expression level, measured on the basis of the transcribed mRNA or of the subsequently translated protein, is quantitatively more than 50%, preferably 100%, particularly preferably 500%, very particularly preferably 1000%, higher than a comparative value obtained with the promoter described by SEQ ID NO: 1, 2 or 3, under otherwise unchanged conditions. The comparative value is preferably the expression level of the natural mRNA of the particular gene or of the natural gene product. A further preferred comparative value is the expression level obtained using a random but particular nucleic acid sequence, preferably those nucleic acid sequences which code for readily quantifiable proteins. In this connection, very particular preference is given to reporter proteins (Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) such as the green fluorescence protein (GFP) (Chui WL et al., Curr Biol 1996, 6:325-330; Leffel SM et al., Biotechniques. 23(5):912-8, 1997), chloramphenicol transferase, a luciferase (Millar et al., Plant Mol Biol Rep 1992 10:324-414) or β-galactosidase, and very particular preference is given to β-glucuronidase (Jefferson et al. (1987) EMBO J. 6:3901-3907).

Otherwise unchanged conditions means the expression initiated by one of the expression cassettes to be compared is not modified by a combination with additional genetic control sequences, for example enhancer sequences. Unchanged conditions further means that all basic conditions such as, for example, plant species, developmental stage of the plants, growing conditions, assay

conditions (such as buffer, temperature, substrates, etc.) are kept identical between the expressions to be compared.

Mutations comprise substitutions, additions, deletions, inversions or insertions of one or more nucleotide residues. Thus, for example, the present invention also includes those nucleotide sequences which are obtained by modification of a promoter of SEQ ID NO: 1, 2 or 3. The aim of such a modification may be the further narrowing down of the sequence comprised therein or else, for example, the introduction of further restriction enzyme cleavage sites, the removal of excess DNA or the addition of further sequences, for example of further regulatory sequences.

Where insertions, deletions or substitutions such as, for example, transitions and transversions are suitable, techniques known per se, such as in vitro mutagenesis, "primer repair", restriction or ligation, may be used. Manipulations such as, for example, restriction, chewing-back or filling-in of protruding ends to give blunt ends can provide complementary fragment ends for ligation. Similar results can be obtained using the polymerase chain reaction (PCR) using specific oligonucleotide primers.

Homology between two nucleic acids means the identity of the nucleic acid sequence over the in each case entire length of the sequence, which is calculated by comparison with the aid of the GAP program algorithm (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), with the parameters set as follows:

Gap Weight: 12	Length Weight: 4
Average Match: 2.912	Average Mismatch:-2.003

By way of example, a sequence which is at least 50% homologous at the nucleic acid level with the sequence SEQ ID NO: 1, 2 or 3 means a sequence which is at least 50% homologous when compared to the sequence SEQ ID NO. 1, 2 or 3 according to the above program algorithm using the above set of parameters.

Functional homologs to the abovementioned promoters for use in the expression cassettes of the invention preferably include those sequences which , are at least 50%, preferably 70%, preferentially at least 80%, particularly preferably at least 90%, very particularly preferably at least 95%, most preferably 99%, homologous over a length of at least 100 base pairs,

preferably at least 200 base pairs, particularly preferably at least 300 base pairs, very particularly preferably at least 400 base pairs and most preferably of at least 500 base pairs.

Further examples of the promoter sequences employed in the expression cassettes or vectors of the invention can readily be found, for example, in various organisms whose genomic sequence is known, such as, for example, *Arabidopsis thaliana*, *Brassica napus*, *Nicotiana tabacum*, *Solanum tuberosum*, *Helianthium anuus*, *Linum sativum* by comparing homologies in databases.

Functional equivalents further means DNA sequences which hybridize under standard conditions with the nucleic acid sequence coding for a promoter according to SEQ ID NO:1, 2 or 3 or with the nucleic acid sequences complementary to it and which have essentially the same properties. Standard hybridization conditions has a broad meaning and means both stringent and less stringent hybridization conditions. Such hybridization conditions are described, inter alia, in Sambrook J, Fritsch EF, Maniatis T et al., in Molecular Cloning - A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, 1989, pp. 9.31-9.57) or in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.

The conditions during the washing step may be selected by way of example from the range of conditions limited by those of low stringency (with approximately 2X SSC at 50°C) and those with high stringency (with approximately 0.2X SSC at 50°C, preferably at 65°C) (20X SSC: 0.3 M sodium citrate, 3 M NaCl, pH 7.0). In addition, the temperature may be raised during the washing step from low stringency conditions at room temperature, approximately 22°C, to higher stringency conditions at approximately 65°C. Both parameters, salt concentration and temperature, may be varied simultaneously, and it is also possible to keep one of the two parameters constant and to vary only the other one. Denaturing agents such as, for example, formamide or SDS may also be employed during hybridization. In the presence of 50% formamide, hybridization is preferably carried out at 42°C. Some exemplary conditions for hybridization and washing are listed below:

(1) Hybridization conditions with, for example,

a) 4X SSC at 65°C, or

b) 6X SSC, 0.5% SDS, 10µg/ml denatured, fragmented salmon sperm DNA at 65°C, or

c) 4X SSC, 50% formamide, at 42°C, or

d) 6X SSC, 0.5% SDS, 10µg/ml denatured, fragmented salmon sperm-DNA, 50% formamide at 42°C, or

e) 2X or 4X SSC at 50°C (low stringency condition), or

f) 2X or 4X SSC, 30 to 40% formamide at 42°C (low stringency condition).

g) 6x SSC at 45°eeeeC, or,

h) 50% formamide, 4xSSC at 42°C, or

i) 50% (vol/vol) formamide, 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer pH 6.5, 75mM NaCl, 75 mM sodium citrate at 42°C, or

j) 0.05 M sodium phosphate buffer pH 7.0, 2 mM EDTA, 1% BSA and 7% SDS.

(2) Washing steps with, for example,

a) 0.1X SSC at 65°C, or

b) 0.1X SSC, 0.5% SDS at 68°C, or

c) 0.1X SSC, 0.5% SDS, 50% formamide at 42°C, or

d) 0.2X SSC, 0.1% SDS at 42°C, or

e) 2X SSC at 65°C (low stringency condition), or

f) 40 mM sodium phosphate buffer pH 7.0, 1% SDS, 2 mM EDTA.

Methods for preparing functional equivalents of the invention preferably comprise introducing mutations into a promoter of SEQ ID NO: 1, 2 or 3. A mutagenesis may be random and the mutagenized sequences are subsequently screened with respect to their properties according to a trial-by-error [sic] procedure. Examples of particularly advantageous selection criteria are an increased resistance to a selection marker and the level of the resulting expression of the introduced nucleic acid sequence.

As an alternative, it is possible to delete non-essential sequences of a promoter of the invention without substantially impairing said properties. Such deletion variants are functionally equivalent fragments of the promoters described by SEQ ID NO: 1, 2 or 3. Examples of such deletion mutants or functionally equivalent fragments, which may be mentioned, are the truncated pFD promoter sequence (pFDs) according to SEQ ID NO: 4 and the truncated TPT promoter sequence according to SEQ ID NO: 27 which, as functionally equivalent parts of their respective source promoters, are expressly included.

The narrowing-down of the promoter sequence to particular essential regulatory regions may also be carried out with the aid of search routines for searching for promoter elements. Particular promoter elements are often present in increased numbers in the regions relevant for promoter activity. Said analysis may be carried out, for example, by computer programs such as the program PLACE ("Plant Cis-acting Regulatory DNA Elements") (K. Higo et al., (1999) Nucleic Acids Research 27:1, 297-300) or the BIOBASE data bank "Transfac" (Biologische Datenbanken GmbH, Brunswick)

Methods for mutagenizing nucleic acid sequences are known to the skilled worker and include, by way of example, the use of oligonucleotides having one or more mutations in comparison with the region to be mutated (for example, within the framework of a site-specific mutagenesis). Typically, primers with from approximately 15 to approximately 75 nucleotides or more are employed, preferably from approx. 10 to approx. 25 or more nucleotide residues being located on both sites of the sequence to be modified. Details and the procedure of said mutagenesis methods are familiar to the skilled worker (Kunkel et al., Methods Enzymol, 154:367-382, 1987; Tomic et al. (1990) Nucl Acids Res 12:1656; Upender, Raj, Weir (1995) Biotechniques 18(1):29-30; US 4,237,224). A mutagenesis may also be carried out by treating, for example, vectors comprising one of the nucleic acid sequences of the invention with mutagenizing agents such as hydroxylamine.

The nucleic acid sequences which are comprised in the expression cassettes of the invention and which are to be expressed transgenically may be functionally linked to further genetic control sequences, in addition to one of the promoters of the invention.

A functional linkage means, for example, the sequential arrangement of a promoter, of the nucleic acid sequence to be expressed transgenically and, where appropriate, of further regulatory elements such as, for example, a terminator in such a way that each of the regulatory elements can carry out its function in the transgenic expression of said nucleic acid sequence, depending on the arrangement of the nucleic acid sequences with respect to sense or antisense RNA. This does not absolutely necessitate a direct linkage in the chemical sense. Genetic control sequences such as, for example, enhancer sequences may exert their function on the target sequence also from relatively distant positions or even from other DNA molecules. Preference is given to arrangements in which the nucleic acid sequence to be expressed transgenically is positioned downstream of the sequence functioning as promoter so that both sequences are covalently linked to one another. The distance between the promoter sequence and the nucleic acid sequence to be expressed transgenically is preferably less than 200 base pairs, particularly preferably less than 100 base pairs and very particularly preferably less than 50 base pairs.

A functional linkage may be prepared by means of common recombination and cloning techniques, as are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987). It is also possible to position further sequences between said two sequences, which have, for example, the function of a linker with particular restriction enzyme cleavage sites or of a signal peptide. Likewise, the insertion of sequences may lead to the expression of fusion proteins.

The term "genetic control sequences" has a broad meaning and means all those sequences which influence the generation or function of the expression cassette of the invention. For example, genetic control sequences modify transcription and translation in prokaryotic or eukaryotic organisms. The expression cassettes of the invention preferably comprise 5'-upstream of the particular nucleic acid sequence to be expressed transgenically one of the promoters of the invention and 3'-downstream a terminator sequence as an additional genetic control sequence and also, where appropriate, further common

regulatory elements which are in each case functionally linked to the nucleic acid sequence to be expressed transgenically.

Genetic control sequences also include further promoters, promoter elements or minimal promoters which may modify the expression-controlling properties. Thus, for example, genetic control sequences can effect tissue-specific expression additionally depending on particular stress factors. Corresponding elements have been described, for example, for water stress, abscisic acid (Lam E and Chua NH (1991) J Biol Chem 266(26): 17131-17135) and heat stress (Schöffl F et al., (1989) Molecular & General Genetics 217(2-3):246-53).

Further promoters which make possible expression in further plant tissues or in other organisms such as, for example, in *E.coli* bacteria may furthermore be functionally linked to the nucleic acid sequence to be expressed. Suitable plant promoters are in principle all of the above-described promoters. It is conceivable, for example, that a particular nucleic acid sequence is transcribed as sense RNA via one promoter (for example one of the promoters of the invention) in one plant tissue and is translated into the corresponding protein, while the same nucleic acid sequence is transcribed to antisense RNA via another promoter having a different specificity in another tissue and the corresponding protein is down-regulated. This may be carried out via an expression cassette of the invention by positioning the one promoter upstream of the nucleic acid sequence to be expressed transgenically and the other promoter downstream of said sequence.

Genetic control sequences furthermore also include the 5'-untranslated region, introns or the noncoding 3'-region of genes, preferably of the pFD, FNR or TPT-genes. It has been demonstrated that these genes may have a substantial function in the regulation of gene expression. Thus it was shown that 5'-untranslated sequences can enhance transient expression of heterologous genes. They may furthermore promote tissue specificity (Rouster J et al.(1998) Plant J. 15:435-440). Conversely, the 5'-untranslated region of the opaque-2 gene suppresses expression. A deletion of the corresponding region leads to an increase in gene activity (Lohmer S et al. (1993) Plant Cell 5:65-73). The nucleic acid sequence indicated under SEQ ID NO:1, 2 or 3 contains the pFD, FNR or TPT-gene section which represents the promoter and the 5'-untranslated region up to the ATG start codon of the respective protein.

McElroy and colleagues (McElroy et al. (1991) Mol Gen Genet 231(1):150-160) reported a construct for transforming monocotyledonous plants, which is based on the rice actin 1 (Act1) promoter. The use of the Act1 intron in combination with the 35S promoter leads in transgenic rice cells to a ten times higher rate of expression compared to the isolated 35S promoter. Optimization of the sequence surrounding the translation initiation site of the reporter gene (GUS) resulted in a four-fold increase of GUS expression in transformed rice cells. A combination of optimized translation initiation site and Act1 intron resulted in a 40-fold increase in GUS expression via the CaMV35S promoter in transformed rice cells; similar results were achieved on the basis of transformed corn cells. Overall, it was concluded from the above-described studies that the expression vectors based on the Act1 promoter are suitable for controlling a sufficiently strong and constitutive expression of foreign DNA in transformed cells of monocotyledonous plants.

The expression cassette may advantageously contain one or more "enhancer sequences" which are functionally linked to the promoter and which enable an increased transgenic expression of the nucleic acid sequence. It is possible to insert additional advantageous sequences such as further regulatory elements or terminators at the 3'-end of the nucleic acid sequences to be expressed transgenically, too. Any of the expression cassettes of the invention may contain one or more copies of the nucleic acid sequences to be expressed transgenically.

Control sequences furthermore means those which enable homologous recombination or insertion into the genome of a host organism or which allow the removal from the genome. In homologous recombination, for example, the natural promoter of a particular gene may be replaced with one of the promoters of the invention. Methods such as the cre/lox technology allow tissue-specific, specifically inducible removal of the expression cassette from the genome of the host organism (Sauer B. (1998) Methods. 14(4):381-92). In this case, particular flanking sequences are attached to the target gene (lox sequences), which later enable a removal by means of the cre recombinase.

The promoter to be introduced may be placed upstream of the target gene to be expressed transgenically by means of homologous recombination by linking the promoter to DNA sequences which are, for example, homologous to endogenous sequences upstream of the reading frame of the target gene. Such sequences are to be understood as genetic control sequences. After a cell has been transformed with the appropriate DNA construct, the two

homologous sequences can interact and thus place the promoter sequence at the desired position upstream of the target gene so that said promoter sequence is now functionally linked to said target gene and forms an expression cassette of the invention. The selection of the homologous sequences determines the insertion point of the promoter. In this case, the expression cassette can be generated by homologous recombination by means of a simple or a doubly-reciprocal recombination. In the case of the singly-reciprocal recombination, only a single recombination sequence is used and the entire introduced DNA is inserted. In the case of the doubly-reciprocal recombination, the DNA to be introduced is flanked by two homologous sequences and the flanked region is inserted. The latter method is suitable for replacing, as described above, the natural promoter of a particular gene with one of the promoters of the invention and thus modifying the location and time of expression of this gene. This functional linkage represents an expression cassette of the invention.

The selection of successfully homologously recombined or else transformed cells normally requires the additional introduction of a selectable marker which imparts to the successfully recombined cells a resistance to a biocide (for example a herbicide), a metabolism inhibitor such as 2-desoxyglucose 6-phosphate (WO 98/45456) or to an antibiotic. The selection marker permits selection of the transformed cells from the untransformed cells (McCormick et al., Plant Cell Reports 5 (1986), 81-84).

Homologous recombination is a relatively rare event in higher eukaryotes, especially in plants. Random integrations into the host genome predominate. One possibility of removing the randomly integrated sequences and thus accumulating cell clones having a correct homologous recombination is the use of a sequence-specific recombination system as described in US 6,110,736. This system consists of three elements: two pairs of specific recombination sequences and a sequence-specific recombinase. This recombinase catalyzes a recombination merely between the two pairs of specific recombination sequences. One pair of these specific DNA sequences is placed outside the DNA sequence to integrated, i.e. outside the two homologous DNA sequences. In the case of a correct homologous recombination, these sequences are not cotransferred into the genome. In the case of a random integration, they normally insert together with the rest of the construct. Using a specific recombinase and a construct comprising a second pair of said specific sequences, the randomly inserted sequences can be excised or inactivated by inversion, while the sequences inserted correctly via homologous

recombination remain in the genome. It is possible to use a multiplicity of sequence-specific recombination systems and the Cre/lox system of bacteriophage P1, the FLP/FRT system of yeast, the Gin recombinase of phase Mu, the *E. coli* Pin recombinase and the R/RS system of the plasmid pSR1 are mentioned by way of example. Preference is given to the bacteriophage P1 Cre/lox and the yeast FLP/FRT system. Here the recombinase (Cre or FLP) interacts specifically with its respective recombination sequences (34bp lox sequence or 47bp FRT sequence) in order to delete or invert the transiently stored sequences. The FLP/FRT and cre/lox recombinase systems have already been applied to plant systems (Odell et al.(1990) Mol. Gen. Genet., 223:369-378)

Polyadenylation signals suitable as control sequences are plant polyadenylation signals and, preferably, those which correspond essentially to *Agrobacterium tumefaciens* T-DNA polyadenylation signals, in particular of the T-DNA gene 3 (octopene synthase) of the Ti plasmid pTiACHS (Gielen et al.,(1984) EMBO J. 3 :(1984), 835 ff) or functional equivalents thereof.

In a particularly preferred embodiment, the expression cassette contains a terminator sequence functional in plants. Terminator sequences functional in plants means in general those sequences which are capable of causing the termination of transcription of a DNA sequence in plants. Examples of suitable terminator sequences are the OCS (octopene synthase) terminator and the NOS (nopaline synthase) terminator. However, particular preference is given to terminator sequences of plants. Terminator sequences of plants means in general those sequences which are part of a natural plant gene. In this connection, particular preference is given to the terminator of the potato cathepsin D inhibitor gene (GenBank Acc. No.: X74985; terminator: SEQ ID NO: 28) or of [sic] the terminator of the field bean storage protein gene VfLE1B3 (GenBank Acc. No.: Z26489; terminator: SEQ ID NO: 29). These terminators are at least equivalent to the viral or T-DNA terminators described in the prior art. The plasmid pSUN5NPTIICat (SEQ ID NO: 24) contains the plant terminator of the potato cathepsin D inhibitor gene.

The skilled worker knows a multiplicity of nucleic acids or proteins whose recombinant expression which is controlled by the expression cassettes or methods of the invention is advantageous. The skilled worker further knows a multiplicity of genes whose repression or elimination by means of expression of a corresponding antisense RNA can likewise achieve advantageous

effects. Advantageous effects which may be mentioned by way of example and not by way of limitation are:

- easier preparation of a transgenic organism, for example by expression of selection markers
- achieving a resistance to abiotic stress factors (heat, cold, drought, increased humidity, environmental toxins, UV radiation)
- achieving a resistance to biotic stress factors (pathogens, viruses, insects and diseases)
- improvement of the properties of food- or feedstuffs
- improvement of growth rate or yield.

Some specific examples of nucleic acids whose expression provides the desired advantageous effects are mentioned below:

1. Selection markers

Selection markers includes both positive selection markers which impart a resistance to an antibiotic, herbicide or biocide and negative selection markers which impart a sensitivity to exactly these substances and also markers which give a growth advantage to the transformed organism (for example by expressing key genes of cytokine biosynthesis; Ebinuma H et al. (2000) Proc Natl Acad Sci USA 94:2117-2121). In the case of positive selection, only those organisms which express the appropriate selection marker grow, while the same organisms die in the case of negative selection. The preparation of transgenic plants prefers the use of a positive selection marker. Furthermore, preference is given to using selection markers which impart growth advantages. Negative selection markers may be used advantageously if particular genes or genome sections are to be removed from an organism (for example in a crossing process).

The selectable marker introduced with the expression cassette imparts to the successfully recombined or transformed cells a resistance to a biocide (for example a herbicide such as phosphinothricin, glyphosate or bromoxynil), a metabolism inhibitor such as 2-desoxyglucose 6-phosphate (WO 98/45456) or to an antibiotic such as, for example, kanamycin, G 418, bleomycin, hygromycin. The selection marker permits selection

of the transformed cells from the untransformed cells (McCormick et al., Plant Cell Reports 5 (1986), 81-84). Particularly preferred selection markers are those which impart a resistance to herbicides. A large number of such selection markers and the sequences coding therefor are known to the skilled worker. Examples which may be mentioned by way of example but not by way of limitation are the following:

i) Positive selection markers:

The selectable marker introduced with the expression cassette imparts to the successfully recombined or transformed cells a resistance to a biocide (for example a herbicide such as phosphinothricin, glyphosate or bromoxynil), a metabolism inhibitor such as 2-desoxyglucose 6-phosphate (WO 98/45456) or to an antibiotic such as, for example, tetracyclines, ampicillin, kanamycin, G418, neomycin, bleomycin or hygromycin. The selection marker permits selection of the transformed cells from the untransformed cells (McCormick et al., Plant Cell Reports 5 (1986), 81-84). Particularly preferred selection markers are those which impart a resistance to herbicides. Examples of selection markers which may be mentioned are:

- DNA sequences coding for phosphinothricin acetyltransferases (PAT) which acetylate the free amino group of the glutamine synthase inhibitor phosphinothricin (PPT) and thus detoxify PPT (de Block et al. 1987, EMBO J. 6, 2513-2518) (also referred to as Bialaphos®-Resistence gene (bar)). The bar gene coding for a phosphinothricin acetyltransferase (PAT) may be isolated, for example, from *Streptomyces hygroscopicus* or *S. viridochromogenes*. Corresponding sequences are known to the skilled worker (from Streptomyces hygroscopicus GenBank Acc. No.: X17220 and X05822, from Streptomyces viridochromogenes GenBank Acc. No.: M 22827 and X65195; US 5,489,520). Furthermore, synthetic genes, for example for expression in plastids, have been described AJ028212 [sic]. A synthetic Pat gene is described in Becker et al. (1994), The Plant J. 5:299-307. Very particular preference is likewise given to the expression of the polypeptide according to SEQ ID NO: 5, for example encoded by a nucleic acid sequence according to SEQ ID NO: 4. The genes impart a resistance to the herbicide Bialaphos® or glufosinate and are frequently used markers in transgenic plants (Vickers , JE et al. (1996). Plant

Mol. Biol. Reporter 14:363–368; Thompson CJ et al. (1987) EMBO Journal 6:2519–2523).

- 5-Enolpyruvylshikimate 3-phosphate synthase genes (EPSP synthasegenes) which impart a resistance to Glyphosat® (N-(phosphonomethyl)glycin). The molecular target of the unselective herbicide glyphosate is 5-enolpyruvyl-3-phosphoshikimate synthase (EPSPS). This enzyme has a key function in the biosynthesis of aromatic amino acids in microbes and plants but not in mammals (Steinrucken HC et al. (1980) Biochem. Biophys. Res. Commun. 94:1207–1212; Levin JG and Sprinson DB (1964) J. Biol. Chem. 239: 1142–1150; Cole DJ (1985) Mode of action of glyphosate; A literature analysis, p. 48–74. In: Grossbard E and Atkinson D (eds.). The herbicide glyphosate. Buttersworths, Boston). Preference is given to using glyphosate-tolerant EPSPS variants as selection markers (Padgette SR et al. (1996). New weed control opportunities: development of soybeans with a Roundup Ready™ gene. In: Herbicide Resistant Crops (Duke, S.O., ed.), pp. 53-84. CRC Press, Boca Raton, FL; Saroha MK and Malik VS (1998) J Plant Biochemistry and Biotechnology 7:65–72). The EPSPS gene of *Agrobacterium sp.* strain CP4 has a natural tolerance for glyphosate, which can be transferred to appropriate transgenic plants. The CP4 EPSPS gene was cloned from *Agrobacterium sp.* strain CP4 (Padgette SR et al.(1995) Crop Science 35(5):1451–1461). 5-Enolpyruvylshikimate 3-phosphate synthases, which are glyphosate-tolerant, as described, for example, in US 5,510,471; US 5,776,760; US 5,864,425; US 5,633,435; US 5,627;061; US 5,463,175; EP 0 218 571, are preferred and the sequences described in each case in the patents have also been deposited with GenBank. Further sequences are described under GenBank Accession X63374. The aroA gene is also preferred (M10947 S. typhimurium aroA locus 5-enolpyruvylshikimate-3-phosphate synthase (aroA protein) gene).

- the gox (glyphosate oxidoreductase) gene coding for the Glyphosat®-degrading enzyme. GOX (for example *Achromobacter sp.* glyphosate oxidoreductase) catalyzes the cleavage of a C-N bond in glyphosate which is thus converted to aminomethylphosphonic acid (AMPA) and glyoxylate. GOX can thereby mediate a resistance to glyphosate (Padgette SR et al. (1996) J Nutr. 1996

Mar;126(3):702-16; Shah D et al. (1986) Science 233: 478-481).

- the deh gene (coding for a dehalogenase which inactivates Dalapon®), (GenBank Acc. No.: AX022822, AX022820 and WO99/27116)

- bxn genes which code for Bromoxynil® -degrading nitrilase enzyme. For example the *Klebsiella ozanenae* nitrilase. Sequences can be found at GenBank, for example under Acc. No: E01313 (DNA encoding bromoxynil-specific nitrilase) and J03196 (K. pneumoniae bromoxynil-specific nitrilase (bxn) gene, complete cds).

- Neomycin phosphotransferases impart a resistance to antibiotics (aminoglycosides) such as neomycin, G418, hygromycin, paromomycin or kanamycin by reducing the inhibiting action thereof by a phosphorylation reaction. Particular preference is given to the nptII gene. Sequences can be obtained from GenBank (AF080390 minitransposon mTn5-GNm; AF080389 minitransposon mTn5-Nm, complete sequence). Moreover, the gene is already part of numerous expression vectors and can be isolated therefrom by using methods familiar to the skilled worker (such as, for example, polymerase chain reaction) (AF234316 pCAMBIA-2301; AF234315 pCAMBIA-2300, AF234314 pCAMBIA-2201). The NPTII gene codes for an aminoglycoside 3'-O-phosphotransferase from *E.coli*, Tn5 (GenBank Acc. No: U00004 position 1401-2300; Beck et al. (1982) Gene 19 327-336).

- the DOG^R1-gene. The DOG^R1 gene was isolated from the yeast *Saccharomyces cerevisiae* (EP 0 807 836). It codes for a 2-desoxyglucose 6-phosphate phosphatase which imparts resistance to 2-DOG (Randez-Gil et al. 1995, Yeast 11, 1233-1240; Sanz et al. (1994) Yeast 10:1195-1202, Sequence: GenBank Acc. No.: NC001140 chromosome VIII, *Saccharomyces cervisiae* position 194799-194056).

- Sulfonylurea- and imidazolinone-inactivating acetolactate synthases which impart a resistance to imidazolinone/sulfonylurea herbicides. Examples of imidazolinone herbicides which may be mentioned are the active substances imazamethabenz-methyl, imazamox, imazapyr, imazaquin, imazethapyr. Examples of sulfonylurea herbicides which may be mentioned are amidosulforon [sic], azimsulfuron, chlorimuronethyl,

chlorsulfuron, cinosulfuron, imazosulforon [sic], oxasulforon [sic], prosulforon [sic], rimsulforon [sic], sulfosulforon [sic]. Numerous further active substances of said classes are known to the skilled worker. An example of a suitable sequence is the sequence of *Arabidopsis thaliana* Csr 1.2 gene deposited under the GenBank Acc-No.: X51514 (EC 4.1.3.18) (Sathasivan K et al. (1990) Nucleic Acids Res. 18(8):2188). Acetolactate synthases which impart a resistance to imidazolinon herbicides are furthermore described under GenBank Acc. No.:

a) AB049823 *Oryza sativa* ALS mRNA for acetolactate synthase, complete cds, herbicide resistant biotype
b) AF094326 *Bassia scoparia* herbicide resistant acetolactate synthase precursor (ALS) gene, complete cds
c) X07645 Tobacco acetolactate synthase gene, ALS SuRB (EC 4.1.3.18)
d) X07644 Tobacco acetolactate synthase gene, ALS SuRA (EC 4.1.3.18)
e) A19547 Synthetic nucleotide mutant acetolactate synthase
f) A19546 Synthetic nucleotide mutant acetolactate synthase
g) A19545 Synthetic nucleotide mutant acetolactate synthase
h) I05376 Sequence 5 from Patent EP 0257993
i) I05373 Sequence 2 from Patent EP 0257993
j) AL133315

Preference is given to expressing an acetolactate synthase according to SEQ ID NO: 7, for example encoded by a nucleic acid sequence according to SEQ ID NO: 6.

- Hygromycin phosphotransferases (X74325 P. pseudomallei gene for hygromycin phosphotransferase) which impart a resistance to the antibiotic hygromycin. The gene is part of numerous expression vectors and can be isolated therefrom by using methods familiar to the skilled worker (such as, for example, polymerase chain reaction) (AF294981 pINDEX4; AF234301 pCAMBIA-1380; AF234300 pCAMBIA-1304; AF234299 pCAMBIA-1303; AF234298 pCAMBIA-1302; AF354046 pCAMBIA-1305.; AF354045 pCAMBIA-1305.1)

- Genes for resistance to

  a) chloramphenicol (chloramphenicol acetyltransferase),

  b) tetracycline, various resistance genes have been described, for example X65876 *S. ordonez* genes class D tetA and tetR for tetracycline resistance and repressor proteins X51366 *Bacillus cereus* plasmid pBC16 tetracycline resistance gene. The gene is also already part of numerous expression vectors and can be isolated therefrom by using methods familiar to the skilled worker (such as, for example, polymerase chain reaction),

  c) streptomycin, various resistance genes have been described, for example under GenBank Acc. No.:AJ278607 Corynebacterium acetoacidophilum ant gene for streptomycin adenylyltransferase.

  d) zeocin, the corresponding resistance gene is part of numerous cloning vectors (e.g. L36849 cloning vector pZEO) and can be isolated therefrom by using methods familiar to the skilled worker (such as, for example, polymerase chain reaction),

  e) ampicillin (β-lactamase gene; Datta N, Richmond MH.(1966) Biochem J. 98(1):204-9; Heffron F et al (1975) J. Bacteriol 122: 250–256; the amp gene was initially cloned for preparing the *E. coli* vectors pBR322; Bolivar F et al. (1977) Gene 2:95–114). The sequence is part of numerous cloning vectors and can be isolated therefrom by using methods familiar to the skilled worker (such as, for example, polymerase chain reaction).

- Genes such as the [sic] isopentenyl transferase from *Agrobacterium tumefaciens* (strain:P022) (Genbank Acc. No.: AB025109). The ipt gene is [lacuna] a key enzyme of cytokine biosynthesis. Its overexpression facilitates the regeneration of plants (e.g. selection of cytokine-free medium). The method for using the ipt gene has been described (Ebinuma H et al. (2000) Proc Natl Acad Sci USA 94:2117-2121; Ebinuma, H et al. (2000) Selection of Marker-free transgenic plants using the oncogenes (*ipt*, *rol* A, B, C) of Agrobacterium as selectable markers, In

Molecular Biology of Woody Plants. Kluwer Academic Publishers).

Various other positive selection markers which impart to the transformed plants a growth advantage over untransformed plants and methods of their use are described, inter alia, in EP-A 0 601 092. Examples which may be mentioned are β-glucuronidase (in connection with, for example, cytokinine glucuronide), mannose 6-phosphate isomerase (in connection with mannose), UDP-galactose 4-epimerase (in connection with, for example, galactose), mannose 6-phosphate isomerase in connection with mannose being particularly preferred.

ii) Negative selection markers

Negative selection markers make possible, for example, the selection of organisms in which sequences comprising the marker gene have been successfully deleted (Koprek T et al. (1999) The Plant Journal 19(6):719-726).In negative selection, for example, a compound which otherwise has no disadvantageous effect on the plant is converted to a compound having a disadvantageous effect, due to the negative selection marker introduced into the plant. Genes which have a disadvantageous effect per se, such as, for example, TK thymidine kinase (TK), and diphtheria toxin A fragment (DT-A), the codA gene product coding for a cytosine deaminase (Gleave AP et al. (1999) Plant Mol Biol. 40(2):223-35; Perera RJ et al. (1993) Plant Mol. Biol 23(4): 793-799; Stougaard J; (1993) Plant J 3:755-761), the cytochrom P450 gene (Koprek et al. (1999) Plant J. 16:719-726), genes coding for a haloalkane dehalogenase (Naested H (1999) Plant J. 18:571-576), the iaaH gene (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) and the tms2 gene (Fedoroff NV & Smith DL 1993, Plant J 3: 273-289) are also suitable.

The concentrations of the antibiotics, herbicides, biocides or toxins, used in each case for selection, have to be adapted to the particular assay conditions or organisms. Examples which may be mentioned for plants are kanamycin (Km) 50 mg/l, hygromycin B 40 mg/l, phosphinothricin (Ppt) 6 mg/l.

It is furthermore possible to express functional analogs of said nucleic acids coding for selection markers. Functional analogs here means all those sequences which have essentially the same function, i.e. which are capable of selection of transformed organisms. In this connection, the functional analog may quite possibly differ in other features. It may

have, for example, a higher or lower activity or else further functionalities.

2. Improved protection of the plant against abiotic stress factors such as drought, heat or cold, for example by overexpression of antifreeze-polypeptides from *Myoxocephalus Scorpius* (WO 00/00512), *Myoxocephalus octodecemspinosus*, of *Arabidopsis thaliana* transcription activator CBF1, of glutamate dehydrogenases (WO 97/12983, WO 98/11240), calcium-dependent protein kinase genes (WO 98/26045), calcineurins (WO 99/05902), farnesyl transferases (WO 99/06580, Pei ZM et al., Science 1998, 282: 287-290), ferritin (Deak M et al., Nature Biotechnology 1999, 17:192-196), oxalate oxidase (WO 99/04013; Dunwell JM Biotechnology and Genetic Engineering Reviews 1998, 15:1-32), DREB1A-Factor (dehydration response element B 1A; Kasuga M et al., Nature Biotechnology 1999, 17:276-286), of genes of mannitol or trehalose synthesis, such as trehalose phosphate synthase or trehalose phosphate phosphatase (WO 97/42326), or by inhibition of genes such as trehalase (WO 97/50561). Particular preference is given to nucleic acids which code for the *Arabidopsis thaliana* transcription activator CBF1 (GenBank Acc. No.: U77378) or for the *Myoxocephalus octodecemspinosus* antifreeze protein (GenBank Acc. No.: AF306348) or functional equivalents of the same.

3. Expression of metabolic enzymes for use in the feed and food sectors, for example expression of phytase and cellulases. Particular preference is given to nucleic acids such as the artificial cDNA coding for a microbial phytase (GenBank Acc. No.: A19451) or functional equivalents thereof.

4. Achieving a resistance, for example to fungi, insects, nematodes and diseases, by specific isolation or accumulation of particular metabolites or proteins in the embryonic epidermis. Examples which may be mentioned are glucosinolates (repulsion of herbivores), chitinases or glucanases and other enzymes which destroy the cell wall of parasites, ribosome-inactivating proteins (RIPs) and other proteins of resistance and stress reactions of the plant, such as those induced by injury or microbial infection of plants or chemically by, for example, salicylic acid, jasmonic acid or ethylene, lysozymes from sources other than plants, such as, for example, T4 lysozyme or lysozyme from various mammals, insecticidal proteins such as *Bacillus thuringiensis* endotoxin, α-amylase inhibitor or protease inhibitors (cowpea trypsine inhibitor), glucanases, lectins such as

phytohemagglutinin, snowdrop lectin, wheat germ agglutinine, RNases and ribozymes. Particular preference is given to nucleic acids coding for chit42 endochitinase from *Trichoderma harzianum* (GenBank Acc. No.: S78423) or for the N-hydroxylating, multifunctional cytochrome P-450 (CYP79) protein from *Sorghum bicolor* (GenBank Acc. No.: U32624) or functional equivalents thereof.

What is known is the accumulation of glucosinolates in plants of the genus of Cardales, in particular of oilseeds, for protection against pests (Rask L et al.(2000) Plant Mol Biol 42:93-113; Menard R et al. (1999) Phytochemistry 52:29-35), the expression of the *Bacillus thuringiensis* endotoxin under the control of the 35 S CaMV promoter (Vaeck et al. (1987) Nature 328:33-37) or the protection of tobacco against fungal infection by expression of a bean chitinase under the control of the CaMV promoter (Broglie et al. (1991) Science 254:1194-1197).

The expression of the snowdrop (*Galanthus nivalis*) lectin agglutinine can achieve a resistance to pests such as the rice pest *Nilaparvata lugens*, for example in transgenic rice plants (Rao et al. (1998) Plant J. 15(4):469-77.). *Nilaparvata lugens* belongs to the phloem-sucking pests and, in addition, acts as a transmitter of important virus-based plant diseases.

The expression of synthetic cryIA(b) and cryIA(c) genes which code for lepidoptera-specific delta-entotoxins from *Bacillus thuringiensis*, can cause a resistance to insect pests in various plants. Thus it is possible to achieve a resistance in rice to two of the most important rice insect pests, the striped stem borer (*Chilo suppressalis*) and the yellow stem borer (*Scirpophaga incertulas*), (Cheng X et al. (1998) Proc Natl Acad Sci USA 95(6):2767-2772; Nayak P et al. (1997) Proc Natl Acad Sci USA 94(6):2111-2116).

5. Expression of genes which cause accumulation of fine chemicals such as tocopherols, tocotrienols or carotenoids. Phytoene desaturase may be mentioned as an example. Preference is given to nucleic acids which code for *Narcissus pseudonarcissus* phytoene desaturase (GenBank Acc. No.: X78815) or functional equivalents thereof.

6. Production of nutraceuticals such as, for example, polyunsaturated fatty acids such as, for example, arachidonic acid or EP (eicosapentenoic acid) or DHA (docosahexaenoic acid) by expressing fatty-acid elongases and/or desaturases or by producing proteins having an improved nutritional value such as, for example, a high proportion of essential amino acids (e.g. the methionine-rich brazil nut albumingen [sic]). Preference is given to nucleic acids coding for the methionine-rich *Bertholletia excelsa* 2S albumin (GenBank Acc. No.:AB044391), the *Physcomitrella patens* Δ6-acyllipid desaturase (GenBank Acc. No.: AJ222980; Girke et al 1998, The Plant Journal 15:39-48), the *Mortierella alpina* Δ6-desaturase (Sakuradani et al 1999 Gene 238:445-453), the *Caenorhabditis elegans* Δ5-desaturase (Michaelson et al. 1998, FEBS Letters 439:215-218), the *Caenorhabditis elegans* Δ5-fatty-acid desaturase (des-5) (GenBank Acc. No.: AF078796), the *Mortierella alpina* Δ5-desaturase (Michaelson et al. JBC 273:19055 - 19059), the *Caenorhabditis elegans* Δ6-elongase (Beaudoin et al. 2000, PNAS 97:6421-6426), the *Physcomitrella patens* Δ6-elongase (Zank et al. 2000, Biochemical Society Transactions 28:654-657) or functional equivalents thereof.

7. Production of fine chemicals (such as, for example, enzymes) and pharmaceuticals (such as, for example, antibodies or vaccines, as described in Hood EE, Jilka JM. (1999) Curr Opin Biotechnol. 10(4):382-6; Ma JK, Vine ND (1999) Curr Top Microbiol Immunol 236:275-92). For example, it was possible to produce on a large scale recombinant avidin from egg white and bacterial β-glucuronidase (GUS) in transgenic corn plants (Hood et al. (1999) Adv Exp Med Biol 464:127-47. Review). These recombinant proteins from corn plants are sold by Sigma (Sigma Chemicals Co.) as high-purity biochemicals.

8. Achieving an increased storage capability in cells which usually contain relatively few storage proteins or storage lipids, with the aim of increasing the yield of said substances, for example by expressing an acetyl-CoA carboxylase. Preference is given to nucleic acids coding for *Medicago sativa* acetyl-CoA carboxylase (accase) (GenBank Acc. No.: L25042) or functional equivalents thereof.

Further examples of advantageous genes are mentioned, for example, in Dunwell JM, Transgenic approaches to crop improvement, J Exp Bot. 2000;51 Spec No; pages 487-96.

It is furthermore possible to express functional analogs of the nucleic acids and proteins mentioned. Functional analogs here means all those sequences which have essentially the same function, i.e. which are capable of the same function (for example substrate conversion or signal transduction) as the protein mentioned by way of example. The functional analog may quite possibly differ in other features. It may have, for example, a higher or lower activity or else have further functionalities. Functional analogs further means sequences which code for fusion proteins comprising one of the preferred proteins and other proteins, for example another preferred protein, or else a signal peptide sequence.

The nucleic acids may be expressed under the control of the promoters of the invention in any desired cell compartment such as, for example, the endomembrane system, the vacuole and the chloroplasts. Desired glycosylation reactions, particular foldings, and the like are possible by utilizing the secretory pathway. Secretion of the target protein to the cell surface or secretion into the culture medium, for example when using suspension-cultured cells or protoplasts, is also possible. The required target sequences may both be taken into account in individual vector variations and be introduced into the vector together with the target gene to be cloned by using a suitable cloning strategy. Target sequences which may be used are both endogenous, if present, and heterologous sequences. Additional heterologous sequences which are preferred for functional linkage but not limited thereto are further targeting sequences for ensuring subcellular localization in the apoplast, in the vacuole, in plastids, in mitochondria, in the endoplasmic reticulum (ER), in the nucleus, in elaioplasts or other compartments; and also translation enhancers such as the 5'-leader sequence from tobacco mosaic virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711) and the like. The method of transporting proteins which are per se not located in the plastids specifically into said plastids has been described, (Klosgen RB und Weil JH (1991) Mol Gen Genet 225(2):297-304; Van Breusegem F et al. (1998) Plant Mol Biol. 38(3):491-496). Preferred sequences are:

a) small subunit (SSU) of ribulose bisphosphate carboxylase (Rubisco ssu) from pea, corn, sunflower

b) transit peptides derived from genes of fatty-acid biosynthesis in plants, such as the transit peptide of the plastid acyl carrier protein (ACP), stearyl-ACP desaturase, β-ketoacyl-ACP synthase or acyl-ACP thioesterase.

c) the transit peptide for GBSSI ("granule bound starch synthase I")

d) LHCP II genes.

The target sequences may be linked to other targeting sequences which differ from the transit peptide-encoding sequences, in order to ensure subcellular localization in the apoplast, in the vacuole, in plastids, in the mitochondrion, in the endoplastic reticulum (ER), in the nucleus, in elaioplasts or in other compartments. It is also possible to use translation enhancers such as the 5'-leader sequence from tobacco mosaic virus (Gallie et al. (1987), Nucl. Acids Res. 15: 8693-8711) and the like.

The skilled worker further knows that there is no need for him to express the above-described genes directly by using the nucleic acid sequences coding for said genes or to repress them by antisense, for example. He may also use, for example, artificial transcription factors of the zinc finger protein type (Beerli RR et al. (2000) Proc Natl Acad Sci USA 97(4):1495-500). These factors attach to the regulatory regions of the endogenous genes to be expressed or repressed and cause expression or repression of the endogenous gene, depending on the design of the factor. Thus it is also possible to achieve the desired effects by expressing an appropriate zinc finger transcription factor under the control of one of the promoters of the invention.

It is likewise possible to use the expression cassettes of the invention for suppressing or reducing the replication or/and translation of target genes by gene silencing.

The expression cassettes of the invention may also be employed for expressing nucleic acids which mediate "antisense" effects and thus are capable of reducing the expression of a target protein, for example.

Preferred genes and proteins whose suppression results in an advantageous phenotype include by way of example but not by way of limitation:

a) polygalacturonase for preventing cell degradation and preventing plants and fruits, for example tomatoes, from becoming "mushy". Preference is given to using for this nucleic acid sequences such as that of the tomato polygalacturonase gene (GenBank Acc. No.: X14074) or its homologs from other genera and species.

b) reducing the expression of allergenic proteins, as described, for example, in Tada Y et al. (1996) FEBS Lett 391(3):341-345 or Nakamura R (1996) Biosci Biotechnol Biochem 60(8):1215-1221.

c) modifying the color of flowers by suppressing the expression of enzymes of anthocyane biosynthesis. Appropriate procedures have been described (for example in Forkmann G, Martens S. (2001) Curr Opin Biotechnol 12(2):155-160). Preference is given to using for this nucleic acid sequences such as those of flavonoid 3'-hydroxylase (GenBank Acc. No.: AB045593), dihydroflavanol 4-reductase (GenBank Acc. No.: AF017451), chalcone isomerase (GenBank Acc. No.: AF276302), chalcone synthase (GenBank Acc. No.: AB061022), flavanone 3-beta-hydroxylase (GenBank Acc. No.: X72592) and flavone synthase II (GenBank Acc. No.: AB045592) and the homologs thereof from other genera and species.

d) altering the amylose/amylopectin content in starch by suppressing the branching enzyme Q which is responsible for the α-1,6-glycosidic linkage. Appropriate procedures have been described (for example in Schwall GP et al. (2000) Nat Biotechnol 18(5):551-554). Preference is given to using for this nucleic acid sequences such as that of the potato starch branching enzyme II (GenBank Acc. No.: AR123356; US 6,169,226) or its homologs from other genera and species.

An antisense nucleic acid first means a nucleic acid sequence which is completely or partially complementary to at least a part of the sense strand of said target protein. The skilled worker knows that it is possible to use, as an alternative, the cDNA or the corresponding gene as starting template for corresponding antisense constructs. Preferably, the antisense nucleic acid is complementary to the coding region of the target protein or to a part thereof. However, the antisense nucleic acid may also be complementary to the noncoding region or to a part thereof. Starting from the sequence information for a target protein, it is possible to design an antisense nucleic acid in the manner familiar to the skilled worker by taking into account the Watson and Crick base pairing rules. An antisense nucleic acid may be

complementary to the entire or to a part of the nucleic acid sequence of a target protein. In a preferred embodiment, the antisense nucleic acid is an oligonucleotide of, for example, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length.

In a preferred embodiment, the antisense nucleic acid comprises α-anomeric nucleic acid molecules. α-anomeric nucleic acid molecules form particular double-stranded hybrids with complementary RNA, in which, in contrast to the normal β-units, the strands run parallel to one another (Gautier et al. (1987) Nucleic Acids. Res. 15:6625-6641).

Likewise included is the use of the above-described sequences in sense orientation, which may lead to cosuppression, as is familiar to the skilled worker. It has been demonstrated in tobacco, tomato and petunia that expression of sense RNA of an endogenous gene can reduce or eliminate expression of said gene, in a similar manner to what has been described for antisense approaches (Goring et al. (1991) Proc. Natl Acad Sci USA, 88:1770-1774; Smith et al. (1990) Mol Gen Genet 224:447-481; Napoli et al. (1990) Plant Cell 2:279-289; Van der Krol et al. (1990) Plant Cell 2:291-299). The introduced construct may represent the gene to be reduced completely or only partially. The possibility of translation is not required.

Very particular preference is also given to the use of methods such as gene regulation by means of double-stranded RNA (double-stranded RNA interference). Relevant methods are known to the skilled worker and have been described in detail (e.g. Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). The processes and methods described in the references listed are hereby expressly incorporated by reference. The simultaneous introduction of strand and complementary strand causes here a highly efficient suppression of native genes.

Advantageously, the antisense strategy may be coupled with a ribozyme method. Ribozymes are catalytically active RNA sequences which, coupled to the antisense sequences, catalytically cleave the target sequences (Tanner NK. FEMS Microbiol Rev. 1999; 23 (3):257-75). This can increase the efficiency of an antisense strategy. The expression of ribozymes in order to reduce particular proteins is known to the skilled worker and is described, for example, in EP-A1 0 291 533, EP-A1 0 321 201 and EP-A1 0 360 257. Suitable target sequences and ribozymes may be determined, for example, as described in Steinecke (Ribozymes,

Methods in Cell Biology 50, Galbraith et al eds Academic Press, Inc. (1995), 449-460), by calculations of the secondary structure of ribozyme RNA and target RNA and by the interaction thereof (Bayley CC et al., Plant Mol Biol. 1992; 18(2):353-361; Lloyd AM and Davis RW et al., Mol Gen Genet. 1994 Mar;242(6):653-657). An example which may be mentioned is hammerhead ribozymes (Haselhoff and Gerlach (1988) Nature 334:585-591). Preferred ribozymes are based on derivatives of Tetrahymena L-19 IVS RNA (US 4,987,071; US 5,116,742). Further ribozymes with selectivity for an L119 mRNA may be selected (Bartel D und Szostak JW (1993) Science 261:1411-1418).

In another embodiment, target protein expression may be reduced using nucleic acid sequences which are complementary to regulatory elements of the target protein genes and which form together with said genes a triple-helical structure and thus prevent gene transcription (Helene C (1991) Anticancer Drug Des. 6(6):569-84; Helene C et al. (1992) Ann NY Acad Sci 660:27-36; Maher LJ (1992) Bioassays 14(12):807-815).

The expression cassette of the invention and the vectors derived therefrom may contain further functional elements.

The term functional element has a broad meaning and means all those elements which influence preparation, propagation or function of the expression cassettes of the invention or of vectors or organisms derived therefrom. Examples which may be mentioned but which are not limiting are:

a) reporter genes which code for readily quantifiable proteins and which ensure, via intrinsic color or enzyme activity, an evaluation of the transformation efficiency and of the location or time of expression. In this connection, very particular preference is given to genes coding for reporter proteins (see also Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) such as

   - green fluorescence protein (GFP) (Chui WL et al., Curr Biol 1996, 6:325-330; Leffel SM et al., Biotechniques. 23(5):912-8, 1997; Sheen et al.(1995) Plant Journal 8(5):777-784; Haseloff et al.(1997) Proc Natl Acad Sci USA 94(6):2122-2127; Reichel et al.(1996) Proc Natl Acad Sci USA 93(12):5888-5893; Tian et al. (1997) Plant Cell Rep 16:267-271; WO 97/41228).

- chloramphenicol transferase (Fromm et al. (1985) Proc. Natl. Acad. Sci. USA 82:5824-5828),

- Luciferase (Millar et al., Plant Mol Biol Rep 1992 10:324-414; Ow et al. (1986) Science, 234:856-859); allows bioluminescence detection.

- β-galactosidase, coding for an enzyme for which various chromogenic substrates are available.

- β-glucuronidase (GUS) (Jefferson et al., EMBO J. 1987, 6, 3901-3907) or the uidA gene which encodes an enzyme for various chromogenic substrates.

- R-locus gene product: protein which regulates production of anthocyanine pigments (red color) in plant tissue and thus makes possible a direct analysis of the promoter activity without the addition of additional auxiliary substances or chromogenic substrates (Dellaporta et al., In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, 11:263-282, 1988).

- β-lactamase (Sutcliffe (1978) Proc Natl Acad Sci USA 75:3737-3741), enzyme for various chromogenic substrates (e.g. PADAC, a chromogenic cephalosporin).

- xylE gene product (Zukowsky et al. (1983) Proc Natl Acad Sci USA 80:1101-1105), catechol dioxygenase which can convert chromogenic catechols.

- alpha-amylase (Ikuta et al. (1990) Bio/technol. 8:241-242).

- tyrosinase (Katz et al.(1983) J.Gen Microbiol 129:2703-2714), enzyme which oxidizes tyrosine to give DOPA and dopaquinone which consequently form the readily detectable melanine.

- aequorin (Prasher et al.(1985) Biochem Biophys Res Commun 126(3):1259-1268), may be used in calcium-sensitive bioluminescence detection.

b) replication origins which ensure a propagation of the expression cassettes or vectors of the invention, for example in *E. coli*. Examples which may be mentioned are ORI (origin of DNA replication), the pBR322 ori or the P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

c) elements, for example border sequences, which enable agrobacteria-mediated transfer into plant cells for transfer and integration into the plant genome, such as, for example, the right or left border of T-DNA or the vir region.

d) multiple cloning regions (MCS) allow and facilitate the insertion of one or more nucleic acid sequences.

Various ways to achieve an expression cassette of the invention are known to the skilled worker. An expression cassette of the invention is prepared, for example, by fusing one of the promoters of the invention (or a functional equivalent or functionally equivalent part according to SEQ ID NO: 1, 2 or 3) or a functional equivalent to a nucleotide sequence to be expressed, where appropriate to a sequence coding for a transit peptide, preferably a chloroplast-specific transit peptide, which is preferably located between the promoter and the particular nucleotide sequence, and also with a terminator or polyadenylation signal. For this purpose, common recombination and cloning techniques as described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al.,(1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience are used.

However, an expression cassette means also those constructs in which the promoter without having been functionally linked beforehand to a nucleic acid sequence to be expressed, is introduced into a host genome, for example, via specific homologous recombination or random insertion and takes over there regulatory control over nucleic acid sequences then functionally linked to it and controls transgenic expression of said nucleic acid sequences. Insertion of the promoter, for example by homologous recombination, upstream of a nucleic acid coding for a particular polypeptide results in an expression cassette of the invention, which controls expression of the particular

polypeptide in the plant. Furthermore, the promoter may also be inserted such that antisense RNA of the nucleic acid coding for a particular polypeptide is expressed. As a result, the expression of said particular polypeptide in plants is down-regulated or eliminated.

Analogously, it is also possible to place a nucleic acid sequence to be expressed transgenically downstream of the endogenous natural promoter, for example by homologous recombination, resulting in an expression cassette of the invention, which controls expression of the nucleic acid sequence to be expressed transgenically in the cotyledons of the plant embryo.

The invention further relates to vectors which contain the above-described expression cassettes. Vectors may be, by way of example, plasmids, cosmids, phages, viruses or else agrobacteria.

The invention also relates to transgenic organisms transformed with at least one expression cassette of the invention or one vector of the invention and also to cells, cell cultures, tissue, parts, such as, for example in the case of plant organisms, leaves, roots, etc., or propagation material derived from such organisms.

Organisms, starting or host organisms mean prokaryotic or eukaryotic organisms such as, for example, microorganisms or plant organisms. Preferred microorganisms are bacteria, yeasts, algae or fungi.

Preferred bacteria are bacteria of the genus *Escherichia*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes* or cyanobacteria for example of the genus *Synechocystis*.

Preference is given especially to microorganisms which are capable of infecting plants and thus transferring the cassettes of the invention. Preferred microorganisms are those of the genus *Agrobacterium* and, in particular of the species *Agrobacterium tumefaciens*.

Preferred yeasts are *Candida*, *Saccharomyces*, *Hansenula* and *Pichia*.

Preferred fungi are *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Fusarium*, *Beauveria* or other fungi described in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) on page 15, Table 6.

Host or starting organisms preferred as transgenic organisms are especially plants. Included within the scope of the invention are all genera and species of the higher and lower plants of the plant kingdom. The mature plants, seeds, shoots and seedlings and also parts, propagation material and cultures, for example cell cultures, derived therefrom are also included. Mature plants means plants at any development stage beyond the seedling. Seedling means a young immature plant in an early development stage.

Annual, perennial, monocotyledonous and dicotyledonous plants are preferred host organisms for preparing transgenic plants. The expression of genes is furthermore advantageous in all ornamental plants, useful or ornamental trees, flowers, cut flowers, shrubs or lawns. Plants which may be mentioned by way of example but not by limitation are angiosperms, bryophytes such as, for example, *Hepaticae* (liverworts) and *Musci* (mosses); pteridophytes such as ferns, horsetail and club mosses; gymnosperms such as conifers, cycades, ginkgo and *Gnetalae;* algae such as *Chlorophyceae, Phaeophpyceae, Rhodophyceae, Myxophyceae, Xanthophyceae, Bacillariophyceae* (diatoms) and *Euglenophyceae.*

Preference is given to plants of the following plant families: *Amaranthaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanacea, Sterculiaceae, Tetragoniacea, Theaceae, Umbelliferae.*

Preferred monocotyledonous plants are in particular selected from the monocotyledonous crop plants, for example of the *Gramineae* family, such as rice, corn, wheat, or other cereal species such as barley, malt, rye, triticale or oats, and also sugar cane and all grass species.

Preferred dicotyledonous plants are in particular selected from the dicotyledonous crop plants, for example

Asteraceae such as sunflower, *Tagetes* or *Calendula* and others,

Compositae, particularly the genus *Lactuca,* in particular the species *sativa* (lettuce), and others,

Cruciferae, particularly the genus *Brassica,* very particularly the species *napus* (oilseed rape), campestris (beet), *oleracea cv Tastie* (cabbage), *oleracea cv* Snowball Y

(cauliflower) und *oleracea cv* Emperor (broccoli), and further cabbage species; and the genus *Arabidopsis*, very particularly the species *thaliana*, and also cress or canola, and others,

Cucurbitaceae such as melon, pumpkin or zucchini, and others,

Leguminosae particularly the genus *Glycine*, very particularly the species max (soyabean), soya and also alfalfa, pea, bean plants or peanut, and others,

Rubiaceae, preferably of the subclass *Lamiidae*, such as, for example, *Coffea arabica* or *Coffea liberica* (coffee bush), and others,

Solanaceae, in particular the genus *Lycopersicon*, very particularly the species *esculentum* (tomato), and the genus *Solanum*, very particularly the species *tuberosum* (potato) and *melongena* (aubergine) and also tobacco or paprika, and others,

Sterculiaceae, preferably of the subclass *Dilleniidae*, such as, for example, *Theobroma cacao* (cacao bush) and others,

Theaceae, preferably of the subclass *Dilleniidae*, such as, for example, *Camellia sinensis* or *Thea sinensis* (tea shrub) and others,

Umbelliferae, preferably the genus *Daucus*, very particularly the species *carota* (carrot), and *Apium* (very particularly the species *graveolens dulce* (celery), and others; and the genus *Capsicum*, very particularly the species *annum* (pepper), and others,

and also linseed, soya, cotton, hemp, flax, cucumber, spinach, carrot, sugarbeet and the various tree, nut and vine species, in particular banana and kiwi fruit.

Also included are ornamental plants, useful and ornamental trees, flowers, cut flowers, shrubs and lawns. Plants which may be mentioned by way of example but not by limitation are angiosperms, bryophytes such as, for example, *Hepaticae* (liverworts) and *Musci* (mosses); pteridophytes such as ferns, horsetail and club mosses; gymnosperms such as conifers, cycades, ginkgo and *Gnetalae*, the *Rosaceae* families, such as rose, *Ericaceae* such as rhododendrons and azaleas, *Euphorbiaceae* such as poinsettias and croton, *Caryophyllaceae* such as pinks, *Solanaceae* such as petunias, *Gesneriaceae* such as African violet,

Balsaminaceae such as catch-me-not, *Orchidaceae* such as orchids, *Iridaceae* such as gladioli, iris, freesia and crocus, *Compositae* such as marigold, *Geraniaceae* such as gerania, *Liliaceae* such as dracaena, *Moraceae* such as ficus, *Araceae* such as sweetheart plant, and others.

Most preference is given to *Arabidopsis thaliana*, *Nicotiana tabacum*, *Tagetes erecta*, *Calendula officinalis* and *Brassica napus* and to all genera and species which are used as food- or feedstuffs, such as the cereal species described, or which are suitable for preparing oils, such as oilseeds (e.g. oilseed rape), nut species, soya, sunflower, pumpkin and peanut.

Plant organisms for the purposes of this invention are furthermore other organisms capable of photosynthetic activity, such as, for example, algae or cyanobacteria, and also mosses.

Preferred algae are green algae such as, for example, algae of the genus *Haematococcus*, *Phaedactylum tricornatum*, *Volvox* or *Dunaliella*.

The preparation of a transformed organism or of a transformed cell requires introducing the appropriate DNA into the appropriate host cell. A multiplicity of methods is available for this process which is referred to as transformation (see also Keown et al. 1990 Methods in Enzymology 185:527-537). Thus, by way of example, the DNA may be introduced directly by microinjection or by bombardment with DNA-coated microparticles. The cell may also be permeabilized chemically, for example using polyethylene glycol, so that the DNA can enter the cell via diffusion. The DNA may also be performed [sic] via protoplast fusion with other DNA-comprising units such as minicells, cells, lysosomes or liposomes. Another suitable method for introducing DNA is electroporation in which the cells are reversibly permeabilized by an electric impulse.

In the case of plants, the methods described for transforming and regenerating plants from plant tissues or plant cells are utilized for transient or stable transformation. Suitable methods are especially protoplast transformation by polyethylene glycol-induced DNA uptake, the biolistic method using the gene gun, the "particle bombardment" method, electroporation, the incubation of dry embryos in DNA-comprising solution and microinjection.

Apart from these "direct" transformation techniques, a transformation may also be carried out by bacterial infection by means of *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. These strains contain a plasmid (Ti or Ri plasmid) which is transferred to the plant after *Agrobacteria* infection. A part of this plasmid, denoted T-DNA (transferred DNA) is integrated into the genome of the plant cell.

The *Agrobacterium*-mediated transformation is best suited to dicotyledonous diploid plant cells, whereas the direct transformation techniques are suitable for any cell type.

An expression cassette of the invention may be introduced advantageously into cells, preferably into plant cells, by using vectors.

In an advantageous embodiment, the expression cassette is introduced by means of plasmid vectors. Preference is given to those vectors which enable a stable integration of the expression cassette into the host genome.

In the case of injection or electroporation of DNA into plant cells, no particular demands on the plasmid used are made. It is possible to use simple plasmids such as those of the pUC series. If complete plants are to be regenerated from the transformed cells, it is necessary for an additional selectable marker gene to be present on the plasmid.

Transformation techniques have been described for various monocotyledonous and dicotyledonous plant organisms. Furthermore, various possible plasmid vectors which normally contain an origin of replication for propagation in *E.coli* and a marker gene for selection of transformed bacteria are available for introducing foreign genes into plants. Examples are pBR322, pUC series, M13mp series, pACYC184 etc.

The expression cassette may be introduced into the vector via a suitable restriction cleavage site. The resultant plasmid is first introduced into *E.coli*. Correctly transformed *E.coli* cells are selected, cultivated and the recombinant plasmid is obtained using methods familiar to the skilled worker. Restriction analysis and sequencing may be used in order to check the cloning step.

Transformed cells, i.e. those which contain the introduced DNA integrated into the DNA of the host cell may be selected from untransformed cells, if a selectable marker is part of the

introduced DNA. A marker may be, by way of example, any gene which is capable of imparting a resistance to antibiotics or herbicides. Transformed cells which express such a marker gene are capable of surviving in the presence of concentrations of an appropriate antibiotic or herbicide, which kill an untransformed wild type. Examples are the bar gene which imparts resistance to the herbicide phosphinothricin (Rathore KS et al., Plant Mol Biol. 1993 Mar;21(5):871-884), the nptII gene which imparts resistance to kanamycin, the hpt gene which imparts resistance to hygromycin and the EPSP gene which imparts resistance to the herbicide glyphosate.

Depending on the method of DNA introduction, further genes may be required on the vector plasmid. If agrobacteria are used, the expression cassette is to be integrated into specific plasmids, either into an intermediate vector (shuttle vector) or a binary vector. If, for example, a Ti or Ri plasmid is to be used for transformation, at least the right border, in most cases, however, the right and the left border, of the Ti or Ri plasmid T-DNA is connected as flanking region with the expression cassette to be introduced. Preference is given to using binary vectors. Binary vectors can replicate both in *E.coli* and in *Agrobacterium*. They normally contain a selection marker gene and a linker or polylinker flanked by the right and left T-DNA border sequences. They may be transformed directly into *Agrobacterium* (Holsters et al.,Mol. Gen. Genet. 163 (1978), 181-187). The selection marker gene permits selection of transformed *Agrobacteria;* an example is the nptII gene which imparts a resistance to kanamycin. The *Agrobacterium* which in this case acts as the host organism should already contain a plasmid with the vir region. This region is required for the transfer of T-DNA into the plant cell. An *Agrobacterium* transformed in this way may be used for transformation of plant cells.

The use of T-DNA for transformation of plant cells has been intensely studied and described (EP 120516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasserdam, Chapter V; Fraley et al., Crit. Rev. Plant. Sci., 4:1-46 and An et al., EMBO J. 4 (1985), 277-287). Various binary vectors are known and partly commercially available, such as, for example, pBIN19 (Clontech Laboratories, Inc. U.S.A.).

The DNA is transferred into the plant cell by coculturing plant explants with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. Starting from infected plant material (e.g. leaf, root or stem parts, but also protoplasts or plant cell suspensions), it is possible to regenerate whole plants by using

a suitable medium which may contain, for example, antibiotics or biocides for selection of transformed cells. The plants obtained may then be screened for the presence of the introduced DNA, in this case the expression cassette of the invention. As soon as the DNA has integrated into the host genome, the corresponding genotype is normally stable and the corresponding insertion is also found again in subsequent generations. Normally, the integrated expression cassette contains a selection marker which imparts to the transformed plant a resistance to a biocide (for example a herbicide), a metabolism inhibitor such as 2-DOG or an antibiotic such as kanamycin, G 418, bleomycin, hygromycin or phosphinothricin etc. The selection marker allows the selection of transformed cells from untransformed cells (McCormick et al.,(1986) Plant Cell Reports 5: 81-84). The plants obtained may be cultivated and crossed in the common manner. Two or more generations should be cultured in order to ensure that the genomic integration is stable and heritable.

The abovementioned methods are described, for example, in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993), pp.128 - 143 and in Potrykus ,(1991) Annu. Rev. Plant Physiol. Plant Molec. Biol. 42: 205 - 225). The construct to be expressed is preferably cloned into a vector which is suitable for transforming *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al.,(1984) Nucl. Acids Res. 12: 8711f.).

As soon as a transformed plant cell has been prepared, it is possible to obtain a complete plant by using methods known to the skilled worker. To this end, callus cultures are used as starting point, by way of example. From these still undifferentiated cell masses, it is possible to induce formation of shoot and root in the known manner. The shoots obtained can be planted out and cultivated.

The efficacy of expression of the nucleic acids to be expressed transgenically can be determined, for example, *in vitro* by shoot meristem propagation using one of the above-described selection methods.

The invention further relates to cells, cell cultures, parts, such as, for example, roots, leaves, etc. in the case of transgenic plant organisms, and transgenic propagation material such as seeds or fruits derived from the above-described transgenic organisms.

Genetically modified plants of the invention, which can be consumed by humans and animals, may also be used, for example directly or after preparation known per se, as foodstuffs or feedstuffs.

The invention further relates to the use of the above-described transgenic organisms of the invention and of the cells, cell cultures, parts, such as, for example, roots, leaves, etc., in the case of transgenic plant organisms, and transgenic propagation material such as seeds or fruits for the production of food- or feedstuffs, pharmaceuticals or fine chemicals.

Preference is further given to a method for the recombinant production of pharmaceuticals or fine chemicals in host organisms, in which a host organism is transformed with one of the above-described expression cassettes or vectors and said expression cassette contains one or more structural genes which code for the fine chemical of interest or catalyze the biosynthesis of the fine chemical of interest, and the transformed host organism is cultivated and the fine chemical of interest is isolated from the cultivation medium. This method is broadly applicable for fine chemicals such as enzymes, vitamins, amino acids, sugars, fatty acids, natural and synthetic flavorings, aromatizing substances and colorants. Particular preference is given to the production of tocopherols and tocotrienols and also carotenoids. Cultivation of the transformed host organisms and isolation from said host organisms or from the cultivation medium are carried out by means of the methods known to the skilled worker. The production of pharmaceuticals such as, for example, antibodies or vaccines is described in Hood EE, Jilka JM (1999). Curr Opin Biotechnol. 10(4):382-6; Ma JK, Vine ND (1999) Curr Top Microbiol Immunol. 236:275-92.

Sequences

1. SEQ ID NO: 1
   Promoter and 5'-untranslated region of the *Arabidopsis thaliana* pFD promoter.

2. SEQ ID NO: 2
   Promoter and 5'-untranslated region of the *Arabidopsis thaliana* FNR promoter.

3. SEQ ID NO: 3
   Promoter and 5'-untranslated region of the *Arabidopsis*

thaliana TPT promoter (2038 bp).

4. SEQ ID NO: 4
Promoter and 5'-untranslated region of the truncated *Arabidopsis thaliana* pFDs promoter

5. SEQ ID NO: 5
Nucleic acids coding for a phosphinothricin acetyltransferase.

6. SEQ ID NO: 6
Amino acid sequence coding for a phosphinothricin acetyltransferase.

7. SEQ ID NO: 7
Nucleic acid coding for an acetolactate synthase.

8. SEQ ID NO: 8
Amino acid sequence coding for an acetolactate synthase.

9. SEQ ID NO: 9 - oligonucleotide primer pWL35
5'-GTC GAC GAA TTC GAG AGA CAG AGA GAC GG-3'

10. SEQ ID NO: 10 - oligonucleotide primer pWL36
5'-GTC GAC GGT ACC GAT TCA AGC TTC ACT GC-3'

11. SEQ ID NO: 11 - oligonucleotide primer pFD1
5'-GAG AAT TCG ATT CAA GCT TCA CTG C-3'

12. SEQ ID NO: 12 - oligonucleotide primer pFD2
5'-CCA TGG GAG AGA CAG AGA GAC G-3'

13. SEQ ID NO: 13 - oligonucleotide primer pFD3
5'-acggatccgagagacagagagacggagacaaaa-3'

14. SEQ ID NO: 14 - oligonucleotide primer pFD5
5'-gcggatccaacactcttaacaccaaatcaaca-3'

15. SEQ ID NO: 15 - oligonucleotide primer L-FNR ara
5'-GTCGACGGATCCGGTTGATCAGAAGAAGAAGAAGAAGATGAACT-3'

16. SEQ ID NO: 16 - oligonucleotide primer R-FNR ara
5'-GTCGACTCTAGATTCATTATTTCGATTTTGATTTCGTGACC -3'

17. SEQ ID NO: 17 - oligonucleotide primer L-TPTara
5'-AAGTCGACGGATCCATAACCAAAAGAACTCTGATCATGTACGTACCCATT-3'

18. SEQ ID NO: 18 - oligonucleotide primer R-TPTara
5'-AGACGTCGACTCTAGATGAAATCGAAATTCAGAGTTTTGATAGTGAGAGC-3'

19. SEQ ID NO: 19 - oligonucleotide primer ubi5
5'-CCAAACCATGGTAAGTTTGTCTAAAGCTTA-3'

20. SEQ ID NO: 20 - oligonucleotide primer ubi3
5'-CGGATCCTTTTGTGTTTCGTCTTCTCTCACG-3'

21. SEQ ID NO: 21 - oligonucleotide primer sqs5
5'-GTCTAGAGGCAAACCACCGAGTGTT-3'

22. SEQ ID NO: 22 - oligonucleotide primer sqs3
5'-CGGTACCTGTTTCCAGAAAATTTTGATTCAG-3'

23. SEQ ID NO: 23
binary plasmid pSUN3 (Sungene GmbH & Co KGaA)

24. SEQ ID NO: 24
binary plasmid pSUN5NPTIICat (Sungene GmbH & Co KGaA)

25. SEQ ID NO: 25
binary plasmid pSUN3PatNos (Sungene GmbH & Co KGaA)

26. SEQ ID NO: 26 - oligonucleotide primer 5-TPTara
5'-AAGTCGACGGATCCTGATAGCTTATACTCAAATTCAACAAGTTAT-3'

27. SEQ ID NO: 27
truncated promoter and 5'-untranslated region of the *Arabidopsis thaliana* TPT-Promoters (1318 bp).

28. SEQ ID NO: 28
nucleic acid sequence of the terminator of the potato cathepsin D inhibitor gene (GenBank Acc. No.: X74985)

29. SEQ ID NO: 29
nucleic acid sequence of the terminator of the field bean storage protein gene VfLE1B3 (GenBank Acc. No.: Z26489).

Description of the figures

1. Figure 1a-c: The TPT and the FNR promoters show a comparable expression pattern in green tissue and in flowers of tobacco and potato. GUS-histochemical stains are formed. The intensity of the GUS blue stain corresponds to the shades of gray displayed. The figures show:

In Figure 1a:

A: Potato leaves with a homogeneous intensive stain over the entire leaf region.

B: Tobacco petioles, intensive blue stain, especially on the edges and in the vascular regions (see arrow)

In Figure 1b:

C: Tobacco stems, intensive blue stain, especially on the edges (see arrow)

D: Tobacco internodia

In Figure 1c:

E: Tobacco flower; blue stain, especially in sepals and petals

2. Figure 2a-b: The TPT promoter and the FNR promoter show a different expression pattern in vegetative and germinative storage tissue of tobacco and potato. While the TPT promoter is active here, the FNR promoter shows no expression. GUS histochemical stains of tobacco seeds and tobacco seedlings and also of potato tubers are shown. However, both promoters exhibit again a comparable activity in seedlings. The intensity of the GUS blue stain corresponds to the shades of gray displayed. The figures show:

In Figure 2a:

A: Tobacco seeds. In the case of the TPT promoter, individual blue stained seeds are visible (see arrow). In the case of the FNR promoter, no stains are detectable.

B: Potato tubers. In the case of the TPT promoter, a homogenous strong blue stain of the potato tuber is visible. In the case of the FNR promoter, only a very weak stain is detectable, if at all.

In Figure 2b:

C: Tobacco seedlings (10 days old). Both promoters show a comparable blue stain (see arrow).

3. Expression cassettes for the expression of kanamycin-resistance (nptII) and phosphinothricin-resistance (pat) markers. Cassette A permits expression of kanamycin

resistance under the TPT or FNR promoter, in addition to a phosphinothricin resistance under the NOS promoter. Cassette B permits expression of phosphinothricin resistance under the TPT or FNR promoter, in addition to kanamycin resistance under the NOS promoter.

LB, RB: left and right border, respectively, of *Agrobacterium* T-DNA
nosP: NOS promoter
pat: nucleic acid sequence coding for phosphinothricin acetyltransferase (pat)
nptII: kanamycin resistance gene (Neomycin phosphotransferase)
nosT: NOS terminator
FNR-P: FNR promoter
TPT-P: TPT promoter

4. Regeneration of transformed tobacco plumulae under kanamycin selection pressure (100 mg/l kanamycin). A: transformation with an FNR promoter - nptII construct. B: transformation with a TPT promoter - nptII construct. A comparable efficient regeneration of transformed tobacco plants was observed.

5. Germination of transformed tobacco plants from transgenic tobacco seeds under phosphinothricin selection pressure (10 mg/l phosphinotricin).

A: transformed with an FNR promoter - pat construct.
B: transformed with a TPT promoter - pat construct.
C: control with untransformed tobacco seeds.
A comparably efficient germination of tobacco plants transformed with the FNR promoter-pat construct and the TPT promoter-pat construct was observed, while untransformed tobacco plants treated in a corresponding manner had no resistance.

Examples

General methods:

The chemical synthesis of oligonucleotides may be carried out in a manner known per se, for example according to the phosphoramidite method (Voet, Voet, 2nd edition, Wiley Press New York, pages 896-897). The cloning steps carried out within the framework of the present invention, such as, for example, restriction cleavages, agarose gel electrophoreses, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and

nylon membranes, ligation of DNA fragments, transformation of E. coli cells, cultivation of bacteria, propagation of phages and sequence analysis of recombinant DNA, are carried out as described in Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6. Recombinant DNA molecules are sequenced according to the method of Sanger (Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-5467), using a laser fluorescence DNA sequencer from ABI.

Example 1: Isolation of genomic DNA from *Arabidopsis thaliana* (CTAB method)

Genomic DNA is isolated from *Arabidopsis thaliana* by grinding approx. 0.25 g of leaf material of young plants in the vegetative state in liquid nitrogen to give a fine powder. The pulverulent plant material is introduced together with 1 ml of 65°C CTAB I buffer (CTAB: hexadecyltrimethylammonium bromide, also called cetyltrimethylammonium bromide; Sigma Cat.-No.: H6269) and 20 µl of β-mercaptoethanol into a prewarmed second mortar and, after complete homogenization, the extract is transferred to a 2 ml Eppendorf vessel and incubated with careful regular mixing at 65°C for 1 h. After cooling to room temperature, the mixture is extracted with 1 ml of chloroform/octanol (24:1, equilibrated by shaking with 1M Tris/HCl, pH8.0) by slowly inverting the vessel and the phases are separated by centrifugation at 8,500 rpm (7,500 x g) and room temperature for 5 min. Subsequently, the aqueous phase is extracted again with 1 ml of chloroform/octanol, centrifuged and carefully mixed with 1/10 volume of CTAB II buffer prewarmed to 65°C by inverting the vessel. 1 ml of chloroform/octanol mixture (see above) is then added with careful agitation to the reaction mixture and the phases are again separated by centrifugation at 8,500 rpm (7,500 x g) and room temperature for 5 min. The aqueous lower phase is transferred to a fresh Eppendorf vessel and the upper organic phase is again centrifuged in a fresh Eppendorf vessel at 8,500 rpm (7,500 x g) and room temperature for 15 min. The aqueous phase resulting herefrom is combined with the aqueous phase of the previous centrifugation step and the entire reaction mixture is then mixed with exactly the same volume of prewarmed CTAB III buffer. This is followed by an incubation at 65°C until the DNA precipitates in flakes. This may continue for up to 1 h or be effected by incubation at 37°C overnight. The sediment resulting from the subsequent centrifugation step (5 min, 2000 rpm (500 x g), 4°C) is admixed with 250 µl of CTAB IV buffer prewarmed to 65°C, and the mixture is incubated at 65°C for at least 30 min or until the sediment has completely dissolved. The DNA is then precipitated by mixing the solution with 2.5 volumes of ice-cold ethanol and

incubating at −20°C for 1 h. As an alternative, the reaction mixture is mixed with 0.6 volumes of isopropanol and, without further incubation, immediately centrifuged at 8,500 rpm (7,500 x g) and 4°C for 15 min. The sedimented DNA is washed twice with in each case 1 ml of 80% strength ice-cold ethanol by inverting the Eppendorf vessel, each washing step being followed by another centrifugation (5 min, 8,500 rpm (7,500 x g), 4°C) and the DNA pellet is then dried in air for approx. 15 min. Finally, the DNA is resuspended in 100 µl of TE comprising 100 µg/ml RNase and the mixture is incubated at room temperature for 30 min. After another incubation phase at 4°C overnight, the DNA solution is homogeneous and can be used for subsequent experiments.

Solution for CTAB:

Solution I (for 200 ml):
100 mM Tris/HCl pH 8.0 (2.42 g)
1.4 M NaCl (16.36 g)
20 mM EDTA (8.0 ml of 0.5 M stock solution)
2 %(w/v) CTAB (4.0 g)

The following is added in each case prior to use: 2% β-mercaptoethanol (20 µl for 1 ml of solution I).

Solution II (for 200 ml):
0.7 M NaCl (8.18 g)
10 %(w/v) CTAB (20 g)

Solution III (for 200 ml):
50 mM Tris/HCl pH 8.0 (1.21 g)
10 mM EDTA (4 ml 0.5 M of 0.5 M stock solution)
1 %(w/v) CTAB (2.0 g)

Solution IV (High-salt TE) (for 200 ml):
10 mM Tris/ HCl pH 8.0 (0.242 g)
0,1 mM EDTA (40 µl of 0.5 M stock solution)
1 M NaCl (11.69 g)

Chloroform/Octanol (24:1) (for 200 ml):
192 ml of chloroform
8 ml of octanol

The mixture is equilibrated by shaking 2x with 1 M Tris/HCl pH 8.0 and stored protected from light.

Example 2: Transformation of tobacco, oilseed rape and potato

Tobacco was transformed via infection with Agrobacterium tumefaciens. [sic] according to the method developed by Horsch (Horsch et al. (1985) Science 227: 1229-1231). All constructs used for transformation were transformed into *Agrobacterium tumefaciens* by using the freeze/thaw method (repeated thawing and freezing). The *Agrobacterium* colonies comprising the desired construct were selected on mannitol/glutamate medium comprising 50 µg/ml kanamycin, 50 µg/ml ampicilin and 25 µg/ml rifampicin.

Tobacco plants (*Nicotiana tabacum* L. cv. *Samsun* NN) were transformed by centrifuging 10 ml of an *Agrobacterium tumefaciens* overnight culture grown under selection, discarding the supernatant and resuspending the bacteria in the same volume of antibiotics-free medium. Leaf disks of sterile plants (approx. 1 cm in diameter) were bathed in this bacteria solution in a sterile Petri dish. The leaf disks were then laid out in Petri dishes on MS medium (Murashige und Skoog (1962) Physiol Plant 15:473ff.) comprising 2% sucrose and 0.8% Bacto agar. After incubation in the dark at 25°C for 2 days, they were transferred to MS medium comprising 100 mg/l kanamycin, 500 mg/l Claforan, 1mg/l benzylaminopurine (BAP), 0.2 mg/l naphthylacetic acid (NAA), 1.6% glucose and 0.8% Bacto agar and cultivation was continued (16 hours light / 8 hours dark). Growing shoots were transferred to hormone-free MS medium comprising 2% sucrose, 250 mg/l Claforan and 0.8% Bacto agar.

Oilseed rape was transformed by means of petiole transformation according to Moloney et al. (Moloney MM, Walker JM & Sharma KK (1989) Plant Cell Reports 8:238-242).

Potatoes (*Solanum tuberosum*) were transformed by infecting leaf disks and internodia of in vitro plants with *Agrobacterium tumefaciens* in liquid Murashige Skoog Medium for 20 minutes and then coculturing them in the dark for 2 d. After coculturing, the explants were cultured on solid MS medium which contains instead of sucrose 1.6% glucose (MG) and which has been supplemented with 5 mg/l NAA, 0.1 mg/l BAP, 250 mg/l Timentin and 30 to 40 mg/l kanamycin (KIM), at 21°C in a 16h light/8h dark rhythm. After this callus phase, the explants were placed on shoot induction medium (SIM). SIM was composed as follows: MG + 2 mg/l Zeatinriboside, 0.02 mg/l NAA, 0.02 mg/l GA3, 250 mg/l Timentin, 30 to 40 mg/l kanamycin. Every two weeks, the explants were transferred to fresh SIM. The developing shoots were rooted on MS medium comprising 2% sucrose and 250 mg/l Timentin and 30 to 40 mg/l kanamycin.

Example 3: Studies on the suitability of the putative ferredoxin (pFD) promoter

a) Cloning of the pFD promoters from *Arabidopsis thaliana*

The putative ferredoxin promoter was amplified from genomic *Arabidopsis thaliana* DNA by means of PCR using the primers pWL35 and pWL36. The primer pWL35 starts with the *Sal*I and *Eco*RI restriction cleavage sites which are located immediately upstream of the coding region of the pFD gene and are highlighted in bold type. The primer pWL36 starts with the *Sal*I and *Asp*718 restriction cleavage sites highlighted in bold type.

Primer pWL35 (SEQ ID NO: 9)
5' **GTC GAC GAA TTC** GAG AGA CAG AGA GAC GG 3'

Primer pWL36 (SEQ ID NO: 10)
5' **GTC GAC GGT ACC** GAT TCA AGC TTC ACT GC 3'

Reaction mixture:

1 µl	Genomic *Arabidopsis* DNA (approx. 250 ng)
0.5 µl	Tth polymerase (2U/µl)
3 µl	$Mg(OAc)_2$ (25mM, final conc. 1.5 mM Mg^{2+})
15.2 µl	3.3 x buffer
4 µl	dNTPs (2.5 mM each, Takara, final concentration: 200 µM each)
24.3 µl	H_2O

PCR conditions:

1	cycle at 95°C for 3 min
10	cycles at 94°C for 10 s, 50°C for 20 s and 72°C for 1 min.
20	cycles at 94°C for 10 s, 65°C for 20 s and 72°C for 1 min.
1	cycle at 72°C for 5 min.

followed by cooling to 4°C until further use.

b) Construction of the pFD promoter-GUS expression cassette

The PCR product of the pFD promoter was cloned into the pCRII vector (Invitrogen) and subsequently isolated by means of the *Sal*I restriction cleavage sites introduced by the pair of primers and purified by gel electrophoresis. For fusion with the GUS gene, the approx. 850 bp pFD promoter fragment was cloned into the *Sal*I-cut binary vector pBI101.2 (Clontech Inc.) and the orientation of the fragment was subsequently verified on the

basis of restriction analyses using the endonucleases *Bgl*II and *Bam*HI. The resulting plasmid pFD::GUS was transformed into tobacco. The tobacco plants generated were denoted pFD:GUS.

c) Construction of the pFD promoter-nptII expression cassette

The putative ferredoxin promoter was amplified from genomic *Arabidopsis thaliana* DNA by means of PCR. The primers were used to add the restriction sites EcoRI and NcoI.

Primer pFD1 (SEQ ID NO: 11)
5' GAG AAT TCG ATT CAA GCT TCA CTG C

Primer pFD2 (SEQ ID NO: 12)
5' CCA TGG GAG AGA CAG AGA GAC G

Reaction mixture:

37.5 µl	H_2O
5.0 µl	10X reaction buffer (final concentration Mg^{2+} 1.5 mM)
4.0 µl	dNTP mix (2.5 mM each)
1.0 µl	Primer pFD1 (10 µM)
1.0 µl	Primer pFD2 (10 µM)
0.5 µl	Taq polymerase (Takara, 2U/µl)
1.0 µl	genomic *Arabidopsis* DNA (approx. 250 ng)

PCR conditions:

1	cycle at 94°C for 3 min
10	cycles at 94°C for 10 s, 48°C for 20 s and 72°C for 1 min.
25	cycles at 94°C for 10 s, 65°C for 20 s and 72°C for 1 min.
1	cycle at 72°C for 5 min.

The PCR product was subcloned into the pCRII plasmid (Invitrogen). The plasmid pCAMBIA 2300 (CAMBIA, GPO Box 3200, Canberra ACT 2601, Australia; GenBank Acc. No: AF234315; Binary vector pCAMBIA-2300, complete sequence; Hajdukiewicz P et al. (1994) Plant Mol Biol 25(6):989-994) was cut with EcoRI/NcoI and the pFD promoter fragment was cloned as EcoRI/NcoI fragment from the pCRII plasmid into this vector. In the process, the 35S promoter was removed from the Cambia vector. The resulting plasmid was referred to as pFD promoter:NPTII and transformed into tobacco.

d) Results of GUS analysis of the transgenic tobacco plants

In the context of histochemical investigations, transgenic pFD::GUS tobacco plants showed strong GUS staining in source leaves and weak GUS staining in the tissues of all flower organs. Strong staining in root tissue was only observed in *in vitro* plants whose roots had been exposed to the illumination. Callus growth was induced on the basis of leaf disks which had been punched out of plants identified as pFD::GUS-positive. The callus tissue and also the plant shoots developing therefrom showed GUS staining whose intensity was comparable to that of the GUS staining of CaMV35S::GUS (in pCambia 1304; CAMBIA, GPO Box 3200, Canberra ACT 2601, Australia; GenBank Acc. No.: AF234300, Binary vector pCAMBIA-1304, complete sequence, Hajdukiewicz P et al. (1994) Plant Mol. Biol. 25(6):989-994) (1994)) transgenic plants. The table listed below (Table 3) summarizes the data of quantifying the GUS activity in the anthers and source leaves of selected transgenic pFD::GUS tobacco plants.

Table 3: Quantification of GUS activity in anthers and source leaves of selected transgenic pFD::GUS tobacco plants

	GUS-Activity (pmol [4MU]/mg[protein]/min)	
pFD :: GUS Plant no.	Anthers	'Source' leaves
pFD5	275	3785
pFD11	174	6202
pFD14	362	2898
pFD15	57	2678

The anthers of the mature flowers display no promoter activity. Said activity is weak in closed flowers.

e) Results of the analysis of kanamycin resistance of the transgenic tobacco plants

In order to study the pFD promoter-assisted imparting of resistance to kanamycin, the pFD promoter:NPTII plasmid was transformed into tobacco. The tobacco plants were selectively regenerated on kanamycin (100 mg/l). The plants regenerated from the developing plumulae comprised kanamycin, demonstrating that the pFD promoter had expressed the NPTII gene and thus made selection possible. The results demonstrate that the isolated nucleic acid sequence has the desired advantageous promoter properties, i.e. it exhibits a promoter activity which is suitable for expressing selection markers effectively and its

activity in the pollen is low. The activity in the anthers is normally less than 10% of the activity in the source leaves.

f) Results of GUS analysis of the transgenic potato plants

The pFD:GUS plasmid (cf. Example 3 b) is transformed into potatoes according to the method described in Example 2.

Result of functional studies: The pFD promoter is strongly expressed in the leaves of the transgenic potato plants analyzed. GUS staining was found to be stronger in the leaves of the potato plants than in the leaves of the tobacco plants described. Weak staining of the flowers and no staining of the tubers indicated low expression in the flowers and no expression in the tubers, respectively.

The data demonstrate that this promoter has no activity in the tubers of potato plants and is suitable for the expression of genes, for example of insecticides, in the leaves and other organs above the ground of plants, whose gene products are unwanted in the storage organs.

g) Preparation of deletion variants of the pFD promoter

A further pFD promoter variant is the deletion pFD-short (pFds). For this purpose, the pFD promoter section from base pairs 137 to 837 was amplified using the following primers:

pFD3 (SEQ ID NO: 13):
5'-acggatccgagagacagagagacggagacaaaa-3'

pFD5 (SEQ ID NO: 14):
5'gcggatccaacactcttaacaccaaatcaaca-3'

Reaction mixture:

37.5 µl	H_2O
5.0 µl	10X reaction buffer ("genomic PCR")
4.0 µl	dNTP mix (2.5 mM each)
2.2 µl	25 mM Mg(OAc)2 (final concentration 1.1 mM)
1.0 µl	Primer pFD3 (10 µM)
1.0 µl	Primer pFD5 (10 µM)
0.5 µl	Pfu-turbo polymerase mix
1.0 µl	Genomic *Arabidopsis* DNA (approx. 250 ng)

PCR conditions:
1 cycle at 95°C for 5 min
25 cycles at 94°C for 30 s, 50°C for 60 s and 72°C for

1 min.
1 cycle at 50°C for 60 s, 72°C for 10 min,
followed by cooling to 4°C until further use

The primers comprised recognition sequences for the restriction enzyme BamHI. After BamHI cleavage, the PCR product was ligated into the plasmid pGUSINT37 (see above) which had likewise been cut with BamHI and had been dephosphorylated. Tobacco leaves were bombarded with the resulting construct pFDsGUSINT by means of Biolistics (BioRad). In this connection, microcarriers (25 µg of Gold, Heraeus 0.3 to 3 µm) were treated with 10 µg of plasmid DNA, 2.5 M $CaCl_2$, and 0.1 M spermidine, washed with alcohol and fired at the leaves which were lying on MS medium under a vacuum of 26 inches and a pressure of 1100 psi. The explants were then incubated in MS medium comprising 2% sucrose for 24 h and then histochemically stained with X-gluc. Blue spots indicated the activity of the promoter.

h) Fusing the pFDs promoter to the NPTII gene

The pFDs promoter is excised as BamHI fragment from pFDsGUSINT and its ends are rendered blunt by means of Klenow-"Fill-In" . The fragment obtained is cloned upstream of the NPTII gene of the EcoRV-cut and dephosphorylated plasmid pSUN5NPTIICat (SEQ ID NO: 24). The plasmid pSUN5NPTII is a derivative of plasmid pSUN3 (SEQ ID NO: 23), which contains, apart from nosP/Pat cassette, also a promoterless NPTII gene. This construct makes it possible to assay promoters on their ability to express NPTII. Selection on phosphinothricin-comprising medium may be carried out in parallel.

The resulting plasmid pSun5FdsNPTII is transformed into tobacco. Regenerated and selected shoots showed that the pFDs promoter allows selection for NPTII.

Example 4: Studies on the suitability of the ferredoxin NADPH oxidoreductase (FNR) promoter

a) Cloning of the FNR promoter from *Arabidopsis thaliana*

The putative promoter region of the FNR gene was amplified from genomic DNA by using the oligonucleotide primers L-FNRara and R-FNRara, bypassing the ATG start codon of the FNR gene and retaining four putative stop codons of the open reading frame located upstream. Using the primers L-FNRara and R-FNRara, the FNR promoter was amplified as a 635 bp fragment corresponding to the section of the clone K2A18.15 from position 69493 to position

70127 (including these two nucleotides) from genomic *Arabidopsis thaliana* DNA by means of PCR. The primer L-FNRara starts with the restriction cleavage sites *Sal*I and *Bam*HI highlighted in bold type and is located upstream of the four stop codons of the gene located upstream of the FNR promoter. The primer R-FNRara starts with the *Sal*I and *Xba*I restriction cleavage sites which are located immediately upstream of the ATG start codon of the FNR gene and are highlighted in bold type.

Primer L-FNR ara (44 mer) (SEQ ID NO: 15):
5' **GTC GAC GGA TCC** GGT TGA TCA GAA GAA GAA GAA GAT GAA CT 3'

Primer R-FNR ara (41 mer) (SEQ ID NO: 16):
5' **GTC GAC TCT AGA** TTC ATT ATT TCG ATT TTG ATT TCG TGA CC 3'

The FNR promoter was amplified using a "touchdown" PCR protocol with the use of the ‚Advantage Genomic Polymerase Mix' (Clontech Laboratories, Inc; Catalogue No. #8418-1). The above-mentioned polymerase mix contains a thermostable DNA polymerase from *Thermus thermophilus* (*Tth* DNA polymerase), mixed with a smaller proportion of Vent proofreading 3'-5' polymerase, and the *Tth* start antibody which makes hot-start PCR possible.

Reaction mixture:

36.8 µl	H_2O
5 µl	10X reaction buffer ("genomic PCR")
1 µl	dNTP mix (10 mM each)
2.2 µl	25 mM $Mg(OAc)_2$ (final concentration 1.1 mM)
1 µl	Primer L-FNR ara (10 µM)
1 µl	Primer R-FNR ara (10 µM)
1 µl	50x polymerase mix
2 µl	Genomic *Arabidopsis* DNA (approx. 500 ng)

PCR conditions:

1	cycle at 94°C for 1 min.
10	cycles at 94°C for 30 s and 70°C for 3 min.
32	cycles at 94°C for 30 s and 65°C for 3 min.
1	cycle at 65°C for 4 min.,

followed by cooling to 4°C until further use.

b) Construction of the FNR promoter-GUS expression cassette

After gel-electrophoretic fractionation and purification from the gel using the Quiagen PCR purification kit, the PCR product of the FNR promoter was cloned into the pCRII vector (Invitrogen) via TA cloning. The promoter fragment was then isolated from the resulting plasmid pATFNR1 by digestion with *XbaI/BamHI* by means of the XbaI and BamHI restriction cleavage sites introduced by the pair of primers and purified by gel electrophoresis. For fusion with the GUS gene, the approx. 600bp FNR promoter fragment was cloned into the *XbaI/BamHI*-digested binary vector pBI101. The correct insertion of the correct fragment in the resulting plasmid pATFNR-Bi was then verified on the basis of a restriction analysis using the endonuclease *EcoRV*. The plasmid pATFNR-Bi was used for transformation of tobacco.

For transformation in oilseed rape, the FNR promoter was cloned as SalI fragment of plasmid pCR_ATFNR into the vector pS3NitGUS cut with SalI and XhoI, thereby replacing the nitrilase promoter.

c) Construction of the FNR promoter-PAT expression cassette

In order to study the FNR promoter-assisted imparting of resistance to phosphinothricin, the FNR promoter was cloned as *SalI* fragment from plasmid pATFNR1 into the *SalI*-cut plasmid pSUN3PatNos (SEQ ID NO: 25) upstream of the phosphinothricin resistance gene.

d) Construction of the FNR promoter-NptII expression cassette

In order to impart resistance to kanamycin, the FNR promoter was cloned as *SalI* fragment into the *XhoI*-cut dephosphorylated plasmid pSUN5NptIICat (Sungene GmbH & Co KGaA, SEQ ID NO: 24) upstream of the NPTII resistance gene. The resulting plasmid is referred to as pS5FNRNptII and was transformed into tobacco and oilseed rape.

e) Results of GUS analysis of the transgenic tobacco plants

Qualitative data:

Transgenic FNR::GUS-Tobacco plants displayed strong GUS expression in all green tissues, especially in source leaves, leaf stalks and internodia, and also in all flower organs of fully developed flowers, (ovary, stigma, sepals and petals) with the exception of pollen which showed no GUS activity; a low staining intensity was detected in anthers. In the first analysis of leaf disks of 80 *in vitro* plants, 70 plants displayed strong GUS staining with low variation in staining intensity between the

individual plants. This was regarded as an indication that the FNR promoter contains an element which provides limited positional effects. In the tissue culture plants, the GUS activity of the FNR::GUS plants was markedly lower than the activity of TPT::GUS plants. Transgenic oilseed rape plants displayed the same staining pattern.

Seed material (F1) of the lines FNR 13, FNR 45 and FNR 28 was analyzed with respect to its GUS activity. It turned out that GUS activity was detected neither in resting seeds nor in growing seedlings (3.5 days after sowing).

In later seedlings stages (6 and 10 days after sowing), strong GUS activity was detected in the cotyledons and in the upper region of the seedling axis, whereas no GUS staining was detected in the roots. In seedlings which had been cultivated in the dark, GUS activity was limited to the cotyledons and was overall lower than in the light-germinated seedlings.

Quantitative analysis of the GUS activity in FNR::GUS transgenic tobacco plants (transformed with plasmid pATFNR-Bi) was analyzed [sic] on the first fully developed leaves of tobacco plants 21 days after transfer from the tissue culture to the greenhouse. The data corresponds to the average of four independent measurements.

Table 4: Quantification of GUS activity in leaf material of selected transgenic FNR::GUS tobacco plants (transformed with plasmid pATFNR-Bi)

FNR::GUS Plant No.	Rank (x-strongest GUS activity among 50 plants)	GUS Activity (pmol 4-MU/mg Protein/min)	Standard deviation
13	1	86491	2974
45	2	41726	1829
14	7	23951	2443
28	9	22148	401
17	10	21557	1157
30	20	13444	744
40	26	11972	1144
25	35	7662	519
35	39	5643	96
21	43	2858	194
C2-(WT)	49	28	4

f) Result of analysis of phosphinothricin resistance of the transgenic tobacco plants

The plasmid pSUN3FNRPat was used for transformation of tobacco by using the *Agrobacterium tumefaciens* strain EHA101, as described under 3. The tobacco plants are selectively regenerated either on phosphinothricin (5 mg/l) or, as a control, on kanamycin (100 mg/l). 97% of explants selected under kanamycin pressure (nosP:NPTII) and 40% of explants selected under phosphinothricin pressure (FNR:Bar) developed plumulae. The plants regenerated under phosphinothricin pressure comprised both the kanamycin and the phosphinothricin gene, demonstrating that the FNR promoter had expressed the phosphinothricin acetyltransferase gene and thus made selection possible. Seeds of the transgenic tobacco plants were laid out on MS medium comprising 10 mg/l phosphinothricin and the rate of germination was determined. In contrast to the control of untransformed tobacco seeds, the seedlings developed normally. The gene of phosphinothricin acetyltransferase, which had been transferred and expressed via the FNR promoter, was detected in the progeny of said lines by means of PCR. The results demonstrated that the isolated nucleic acid sequence has the desired advantageous promoter properties, i.e. it shows a promoter activity which is suitable for expressing selection markers effectively and has no activity in pollen.

g) Results of the analysis of kanamycin resistance of the transgenic tobacco and oilseed rape plants

In order to study the FNR promoter-assisted imparting of resistance to kanamycin, the FNR promoter was combined with the NptII gene. The resulting construct pS5FNRNptII was transformed into the *Agrobacterium tumefaciens* strain GV3101[mp90] for transformation in tobacco and oilseed rape.

Seeds of the transgenic tobacco plants were laid out on MS medium comprising 100 mg/l kanamycin and the rate of germination was determined. In contrast to the control of untransformed tobacco seeds, the seedlings developed normally. The gene of neomycin phosphotransferase (nptII), which had been transferred and expressed via the FNR promoter, was detected in the progeny of said lines by means of PCR.

The resulting strains have been used for transformation, as described under Example 2. Selective regeneration was achieved in the presence of 100 mg/l (or 18 mg/l in the case of oilseed rape) kanamycin. The plants regenerated under kanamycin pressure

comprised both the kanamycin and the phosphinothricin gene, demonstrating that the FNR promoter had expressed the NptII gene and thus made selection of the plants possible.

The results demonstrated that the isolated nucleic acid sequence has the desired advantageous promoter properties, i.e. it shows a promoter activity which is suitable for expressing selection markers effectively and has no activity in pollen.

h) Results of GUS analysis of the transgenic potato plants

The analysis of putatively transgenic potato plants showed in 20 lines a strong GUS staining in the leaves, comparable to the expression pattern of tobacco plants. With the exception of 5 plants which showed a very weak staining in the potato tubers, no FNR promoter expression was detected in the remaining plants.

The data demonstrate that this promoter has very weak, if any, activity in the storage organs of potato plants and is suitable for the expression of genes, for example of insecticides, in the leaves and other organs above the ground of plants, whose gene products are unwanted in the storage organs.

Example 5: Studies on the suitability of the triose phosphate translocator (TPT) promoter

a) Cloning of the TPT promoter from *Arabidopsis thaliana*

The putative promoter region of the TPT gene from *Arabidopsis thaliana* was isolated by amplification using the oligonucleotide primers L-TPTara and R-TPTara, the ATG start codon of the TPT gene being bypassed. Using the primers L-TPTara and R-TPTara, the TPT promoter was amplified as a 2038 bp fragment from genomic *Arabidopsis thaliana* DNA by means of PCR (SEQ ID NO: 3). The primer L-TPTara starts with the *Sal*I and *Bam*HI restriction cleavage sites highlighted in bold type. The primer R-TPTara starts with the *Aat*II, *Sal*I and *Xba*I restriction cleavage sites which are located immediately upstream of the ATG start codon of the TPT gene and are highlighted in bold type.

Primer L-TPTara (SEQ ID NO: 17):
5' AAG **TCG ACG GAT** CCA TAA CCA AAA GAA CTC TGA TCA TGT ACG TAC CCA TT 3'

Primer R-TPTara (SEQ ID NO: 18):

5' AGA CGT CGA CTC TAG ATG AAA TCG AAA TTC AGA GTT TTG ATA GTG AGA GC 3'

The TPT promoter was amplified using a "touchdown" PCR protocol with the use of the ‚Advantage Genomic Polymerase Mix' (Clontech Laboratories, Inc; Catalogue No. #8418-1). The above-mentioned polymerase mix contains a thermostable DNA polymerase from *Thermus thermophilus* (*Tth* DNA polymerase), mixed with a smaller proportion of Vent proofreading 3'-5' polymerase, and the *Tth* start antibody which makes hot-start PCR possible.

Reaction mixture:

36.8 µl	H_2O
5 µl	10X reaction buffer ("genomic PCR")
1 µl	dNTP mix (10 mM each)
2.2 µl	25 mM $Mg(OAc)_2$ (final concentration 1.1 mM)
1 µl	Primer L-FNR ara (10 µM)
1 µl	Primer R-FNR ara (10 µM)
1 µl	50x polymerase mix
2 µl	Genomic *Arabidopsis* DNA (approx. 500 ng)

PCR conditions:

1	cycle at 94°C for 1 min.
10	cycles at 94°C for 30 s and 70°C for 3 min.
32	cycles at 94°C for 30 s and 65°C for 3 min.
1	cycle at 65°C for 4 min.,

followed by cooling to 4°C until further use.

b) Construction of the TPT promoter-GUS expression cassette

After gel-electrophoretic fractionation and purification from the gel using the Quiagen PCR purification kit, the PCR product of the TPT promoter was cloned into the pCRII vector (Invitrogen) via TA cloning. The promoter fragment was then isolated from the resulting plasmid pATTPT by means of the *SalI* and *XbaI* restriction cleavage sites introduced by the pair of primers and purified by gel electrophoresis. For fusion with the GUS gene, the approx. 2.0 kb TPT promoter fragment was cloned into the *SalI*/*XbaI*-digested binary vector pBI101. The correct insertion of the correct fragment in the resulting plasmid pATTPT-Bi was then verified on the basis of a restriction analysis using the endonuclease *Hind*III. The plasmid pATTPT-Bi was used for transformation of tobacco.

For transformation in oilseed rape, the TPT promoter was cloned as SalI fragment of plasmid pATTPT into the vector pS3NitGUS cut with SalI and XhoI, thereby replacing the nitrilase promoter.

c) Construction of the TPT promoter-PAT expression cassette

In order to study the TPT promoter-assisted imparting of resistance to phosphinothricin, the TPT promoter was cloned as *Sal*I fragment from plasmid pATTPT into the *Sal*I-cut plasmid pSUN3PatNos upstream of the phosphinothricin resistance gene. The resulting plasmid pSUN3TPTPat was used for transformation of tobacco using the *Agrobacterium tumefaciens* strain EHA101. The tobacco plants were selectively regenerated either on phosphinothricin (5 mg/l) or, as a control, on kanamycin (100 mg/l).

d) Construction of the TPT promoter-NptII expression cassette

In order to impart resistance to kanamycin, the TPT promoter was cloned as *Sal*I fragment into the *Xho*I-cut dephosphorylated plasmid pSUN5NptIICat (Sungene GmbH & Co KGaA, SEQ ID NO: 24) upstream of the NPTII resistance gene. The resulting plasmid is referred to as pS5TPTNptII and was transformed into tobacco and oilseed rape.

e) Results of GUS analysis of the transgenic tobacco plants

Qualitative data

Transgenic TPT::GUS-tobacco plants displayed strong GUS expression in all green tissues, especially in source leaves, here in particular in the trichomes and the flower organs of young and fully developed flowers. GUS activity in the flower region was strongest in the ovaries and in the stigma; staining of the sepals and petals was somewhat weaker. The GUS activity was lowest in the anthers. No GUS activity was detected in the pollen. Transgenic oilseed rape plants showed the same staining pattern.

In the first analysis of leaf disks of 80 *in vitro*-plants, 22 plants showed no staining whatsoever and 22 plants showed strong GUS staining after staining for only 3 hours. The remaining plants displayed a good variety of GUS stainings of various intensities in the individual plants. In the tissue culture plants, the GUS activity of the TPT::GUS plants was markedly stronger than that of the FNR::GUS plants. Seed material (F1) of the lines TPT 55 and TPT 60 were analyzed with respect to their

GUS activity. It turned out that strong, GUS activity was detected both in resting seeds and in growing seedlings (3.5 days after sowing). In later seedling stages (6 and 10 days after sowing), the strongest GUS activity was detected in cotyledons and in the upper region of the seedling axis and a weaker GUS staining in the roots. Seedlings which had been cultivated in the dark displayed an unchanged GUS staining pattern.

Quantitative data

Quantitative analysis of the GUS activity in TPT::GUS transgenic tobacco plants (transformed with plasmid pAT-TPT-Bi) was carried out on the first fully developed leaves of tobacco plants 19 days after transfer from the tissue culture to the greenhouse. The data correspond to the average of four independent measurements.

Table 5: Quantification of GUS activity in leaf tissue of selected transgenic TPT::GUS tobacco plants (transformed with plasmid pAT-TPT-Bi). WT2: controls from untransformed wild-type plants.

TPT::GUS Plant no.	Rank (x-strongest GUS activity among 50 plants)	GUS Activity (pmol 4-MU/mg Protein/min)	Standard deviation
55	1st	62910	3576
15	2nd	58251	2533
10	5th	36759	1008
60	10th	19536	1783
56	11th	18876	1177
43	12th	18858	1404
27	35th	7390	233
44	59th	311	24
80-WT2	80th	5	13

f) Results of GUS analysis of the transgenic potato plants

The analysis of putatively transgenic potato plants showed in 28 lines strong GUS staining in the leaves, comparable to the expression pattern of tobacco plants. A strong staining was likewise detected in the potato tubers of the transgenic plants. This demonstrates that the TPT promoter is expressed strongly and ubiquitously in potatoes, too.

g) Results of the analysis of phosphinothricin resistance of the transgenic tobacco plants

The plasmid pSUN3TPTPat was used for transformation of tobacco by using the *Agrobacterium tumefaciens* strain EHA101, as described under 3. The tobacco plants are selectively regenerated either on phosphinothricin (5 mg/l) or, as a control, on kanamycin (100 mg/l). 97% of explants selected under kanamycin pressure and 70% of explants selected under phosphinothricin pressure developed plumulae. The plants regenerated under phosphinothricin pressure comprised both the kanamycin and the phosphinothricin gene, demonstrating that the TPT promoter had expressed the phosphinothricin acetyltransferase gene and thus made selection possible. Seeds of the transgenic tobacco plants were laid out on MS medium comprising 10 mg/l phosphinothricin and the rate of germination was determined. In contrast to the control of untransformed tobacco seeds, the seedlings developed normally. The gene of phosphinothricin acetyltransferase, which had been transferred and expressed via the TPT promoter, was detected in the progeny of said lines by means of PCR. The results demonstrate that the isolated nucleic acid sequence has the desired advantageous promoter properties, i.e. it shows a promoter activity which is suitable for expressing selection markers effectively and has no activity in pollen.

h) Results of the analysis of kanamycin resistance of the transgenic tobacco and oilseed rape plants

In order to study the TPT promoter-assisted imparting of resistance to kanamycin, the TPT promoter was combined with the NptII gene. The resulting construct pS5TPTNptII was transformed into the *Agrobacterium tumefaciens* strain GV3101[mp90] for transformation in tobacco and oilseed rape.

The resulting strains have been used for transformation, as described under Example 2. Selective regeneration was achieved in the presence of 100 mg/l (or 18 mg/l in the case of oilseed rape) kanamycin. The plants regenerated under kanamycin pressure comprised both the kanamycin and the phosphinothricin gene, demonstrating that the TPT promoter had expressed the NptII gene and thus made selection of the plants possible.

Seeds of the transgenic tobacco plants were laid out on MS medium comprising 100 mg/l kanamycin and the rate of germination was determined. In contrast to the control of untransformed tobacco seeds, the seedlings developed normally. The gene of neomycin

phosphotransferase (nptII), which had been transferred and expressed via the TPT promoter, was detected in the progeny of said lines by means of PCR.

The results demonstrate that the isolated nucleic acid sequence has the desired advantageous promoter properties, i.e. it shows a promoter activity which is suitable for expressing selection markers effectively and has no activity in pollen.

i) Cloning of the truncated TPT promoter (STPT)

The truncated putative promoter region of the *Arabidopsis thaliana* TPT gene (STPT) was isolated from the plasmid pATTPT (SEQ ID NO: 27) by amplification by means of PCR using the primer 5-TPTara (SEQ-ID NO: 26) and R-TPTara (see above SEQ ID NO: 18).

Reaction mixture:

37.8 µl	H_2O
5 µl	10X Reaction buffer ("genomic PCR")
1 µl	dNTP mix (2.5 mM each)
2.2 µl	25 mM $Mg(OAc)_2$ (final concentration 1.1 mM)
1 µl	Primer 5-TPTara (10 µM)
1 µl	Primer R-TPTara (10 µM)
1 µl	50x polymerase mix ("Advantage Genomic Polymerase Mix"; Clontech Lab., Inc.; Cat.-No.: #8418-1)
1 µl	pATTPT plasmid DNA (1 ng)

PCR conditions:

1	cycle at 94°C for 5 min
25	cycles at 94°C for 30 s and 52°C for 1 min.
1	cycle at 52°C for 4 min., followed by cooling to 4°C until further use.

Primer 5-TPTara (SEQ-ID NO: 26)
5'-AAG TCG ACG GAT CCT-GAT-AGC-TTA-TAC-TCA-AAT-TCA-ACA-AGT-TAT-3'

The 1.3 kb PCR product of the truncated TPT promoter was cloned, after gel-electrophoretic fractionation and purification from the gel, into the SmaI-cut and dephosphorylated vector pUC18, using the SureClone Ligation Kit (Amersham Pharmacia Biotech; Code-No.: 27-9300-01). The resulting plasmid is referred to as pATSTPT. The sequence was checked by means of sequencing.

j) Construction of the STPT promoter-NptII expression cassette

In order to impart resistance to kanamycin, the STPT promoter (SEQ ID NO: 27) was cloned as *Sal*I fragment into the *Xho*I-cut dephosphorylated plasmid pSUN5NptIICat (Sungene GmbH & Co KGaA, SEQ ID NO: 24) upstream of the NptII resistance gene. The resulting plasmid is referred to as pS5STPTNptII and was transformed into tobacco and oilseed rape.

k) Results of the analysis of kanamycin resistance of the transgenic tobacco and oilseed rape plants

In order to study the STPT promoter-assisted imparting of resistance to kanamycin, the plasmid pS5STPTNptII was transformed into the *Agrobacterium tumefaciens* strain GV3101[mp90] for transformation into tobacco and oilseed rape. The resulting strain has been used for transformation, as described under Example 2. Selective regeneration was achieved in the presence of 100 mg/l (or 18 mg/l for oilseed rape) kanamycin.

The results demonstrate that the isolated nucleic acid sequence has the desired advantageous promoter properties, i.e. it shows a promoter activity which is suitable for expressing selection markers effectively.

Example 6: Comparison of the transformation efficiencies of the FNR and TPT promoters and of the NOS promoter

In a comparative experiment, the efficiency of the transformation of tobacco was determined using the FNR promoter (FNR-P), the TPT promoter (TPT-P) and the Nos promoter (Nos-P). The promoters were, as described, fused in each case to the NptII gene. After the plumulae had formed and, respectively, the shoot had roots on kanamycin-comprising medium, the resistant transformants were counted and their numbers were compared. A PCR which was used to detect the NptII gene showed the high proportion of transgenic plants.

	NOS-P	FNR-P	TPT-P
Shoot formation	100 %	68 %	76 %
Rooted plants	80 %	81 %	80 %
Transgenic plants	92 %	100 %	100 %

Table 6: Transformation efficiency

Comparative Example 1: Studies of the suitability of the ubiquitin promoter

a) Cloning of the ubiquitin promoter from *Arabidopsis thaliana*

The ubiquitin promoter was amplified from genomic *Arabidopsis thaliana* DNA by means of PCR using the primers ubi5 and ubi3.

ubi5 (SEQ ID NO: 19) :
5'-CCAAACCATGGTAAGTTTGTCTAAAGCTTA-3'

ubi3 (SEQ ID NO: 20):
5'-CGGATCCTTTTGTGTTTCGTCTTCTCTCACG-3'

Reaction mixture:

37.5 µl	H_2O
5 µl	10X reaction buffer ("genomic PCR")
4 µl	dNTP mix (2.5 mM each)
2.2 µl	25 mM $Mg(OAc)_2$ (final concentration 1.1 mM)
1 µl	Primer ubi3 (10 µM)
1 µl	Primer ubi5 (10 µM)
0.5 µl	Pfu-turbo polymerase mix
1 µl	Genomic *Arabidopsis* DNA (ca. 250 ng)

PCR conditions:

1	cycle at 94°C for 5 min.
25	cycles at 94°C for 30 s, 52°C for 1 min. and 72°C for 1 min.
1	cycle at 52°C for 1 min. and 72°C for 10 min., followed by cooling to 4°C until further use.

The resultant PCR fragment was cooled as *Hin*dIII/*Bam*HI fragment into the *Hin*dIII/*Bam*HI-cut plasmid pGUSINT37 (pUBI42GUS) and verified by means of sequence analysis.

b) Cloning of the ubiquitin promoter upstream of the PAT gene

In order to study the ubiquitin promoter-assisted imparting of resistance to phosphinothricin, the ubiquitin promoter was cloned as *Bam*HI/*Hin*dIII fragment into the *Bam*HI/*Hin*dIII-cut plasmid pSUN3PatNos upstream of the phosphinothricin resistance gene. The resulting plasmid pSUN3UBIPat was used for transformation of tobacco using the *Agrobacterium tumefaciens* strain EHA101. The tobacco plants were selectively regenerated either on

phosphinothricin (5 mg/l) or, as a control, on kanamycin (100 mg/l).

c) Results of the analysis of phosphinothricin resistance of the transgenic tobacco plants

In contrast to selection on kanamycin, which was normal, no calli or shoots were obtained under selection on phosphinothricin. Thus, the ubiquitin promoter is unsuitable for expression of a selective marker for the *Agrobacterium tumefaciens*-mediated gene transfer with subsequent regeneration of tissues.

Comparative Example 2: Studies on the suitability of the squalene synthase (SQS) promoter

a) Cloning of the squalene synthase (SQS) promoter from *Arabidopsis thaliana*

The squalene synthase promoter was amplified from genomic *Arabidopsis thaliana* DNA by means of PCR using the primers sqs5 and sqs3.

sqs5 (SEQ ID NO: 21):
5'-GTCTAGAGGCAAACCACCGAGTGTT-3'

sqs3 (SEQ ID NO: 22):
5'-CGGTACCTGTTTCCAGAAAATTTTGATTCAG-3'

Reaction mixture:

37.5 µl	H_2O
5 µl	10X Reaction buffer ("genomic PCR")
4 µl	dNTP mix (2.5 mM each)
2.2 ml	25 mM $Mg(OAc)_2$ (final concentration 1.1 mM)
1 µl	Primer sqs3 (10µM) (10 µM)
1 µl	Primer sqs5 (10 µM)
0.5 µl	Pfu-turbo polymerase mix
1 µl	Genomic *Arabidopsis* DNA (approx. 250 ng)

PCR conditions:

1	cycle at 94°C for 5 min.
25	cycles at 94°C for 30 s, 52°C for 1 min. and 72°C for 1 min.
1	cycle at 52°C for 1 min. and 72°C for 10 min., followed by cooling to 4°C until further use.

The resultant PCR fragment was cloned as *Xba*II/*Bam*HI fragment into the *Xba*II/*Bam*HI-cut plasmid pGUSINT37 (pSQSPGUS) and verified by means of sequence analysis.

b) Cloning of the squalene synthase promoter upstream of the PAT gene

In order to study the squalene synthase promoter-assisted imparting of resistance to phosphinothricin, the squalene synthase promoter was cloned as *Bam*HI/*Sal*I fragment into the *Bam*HI/*Sal*I-cut plasmid pSUN3PatNos upstream of the phosphinothricin resistance gene. The resulting plasmid pSUN3SQSPat was used for transformation of tobacco using the *Agrobacterium tumefaciens* strain EHA101. The tobacco plants were selectively regenerated either on phosphinothricin (5 mg/l) or, as a control, on kanamycin (100 mg/l).

c) Results of the analysis of phosphinothricin resistance of the transgenic tobacco plants

In contrast to selection on kanamycin, which was normal, no calli or shoots were obtained under selection on phosphinothricin. Thus, the ubiquitin [sic] promoter is unsuitable for expression of a selective marker for the *Agrobacterium tumefaciens*-mediated gene transfer with subsequent regeneration of tissues.

Comparative Example 3: Promoter activity assay of the ubiquitin and squalene synthase-promoter by means of a particle gun

In order to assay the activity of the ubiquitin promoter and the squalene synthase promoter, sterile tobacco leaves were bombarded with plasmid DNA of plasmids pUBI42GUS, pSQSPGUS and pGUSINT37 by means of the BioRad Biolistics particle gun. In this connection, microcarriers (25 µg of Gold, Heraeus 0.3 to 3 µm) were treated with 10 µg of plasmid DNA, 2.5 M $CaCl_2$, and 0.1 M spermidine, washed with alcohol and fired at the leaves which were lying on MS agar medium under a vacuum of 26 inches and a pressure of 1100 psi. The explants were then incubated in MS medium comprising 2% sucrose for 24 h and then histochemically stained with X-gluc.

In contrast to the comparative construct pGUSINT37 in which the GUS gene was expressed under the control of the 35S promoter, the ubiquitin promoter and the squalene synthase promoter showed only

very few and very weak GUS-stained dots. This indicates that the ubiquitin and squalene synthase promoter activities are distinctly weaker than the CaMV35S promoter activity.

0817/00024

SEQUENCE LISTING

<110> SunGene GmbH & Co KGaA

<120> Expression cassette for transgenic expression of nucleic acids

<130> NAE3614/2001

<140>
<141>

<160> 29

<170> PatentIn Ver. 2.1

<210> 1
<211> 836
<212> DNA
<213> *Arabidopsis thaliana*

<220>
<221> promoter
<222> (1)..(836)
<223> pFD promoter

<400> 1
gattcaagct tcactgctta aattcacaaa aagagaaaag taagaccaaa ggaataaatc 60
atcctcaaac caaaaacaca tcatacaaaa tcatcaaaca taaatctcca gatgtatgag 120
caccaatcca gttatacaac actcttaaca ccaaatcaac agatttaaca gcgaaataag 180
cttaagccca tacaattatc cgatccaaac aaatataatc gaaaccggca gaggaataag 240
caagtgaatc aaaaagtatg ggacgaggaa gaagatgata cctgaatgag aaagtcaata 300
accttgaccc gaatcgtttt gaagaaaatg gagaaaatcg gttgtatgga ataaaatctt 360
cgaatgatga gatatatgat ctctttggtg tcagtcacat ggcacacgct atcaatttag 420
aaaaacgcgg tggttggtca ccagaattac tacttctcgg tctgatttgg tcatatccgt 480
attaagtccg gttaatattt tccataactg gggtttgaac attcggtttc ttttttcag 540
ttagtccgat ttggagtttt gagtatggaa aaataatact gaatttattt gttcaaactg 600
ttttggaaaa aatatttccc ttaattacga atataattaa aattttaaaa ttcattttat 660
tagatcttgg ttaattcggt ttaatgcatt aatgaatttc ggtttaagtc ggttttcggt 720
ttttatgtcc caccactatc tacaaccgat gatcaacctt atctccgtat tcaccacaaa 780
cagtcatcac tctcacttga cacaaaaact cttttgtctc cgtctctctg tctctc 836

<210> 2
<211> 635
<212> DNA
<213> *Arabidopsis thaliana*

<220>
<221> promoter
<222> (1)..(635)
<223> FNR promoter

<400> 2
ttcattattt cgattttgat ttcgtgacca gcgaacgcag aataccttgt tgtgtaatac 60

tttacccgtg taaatcaaaa acaaaaaggc ttttgagctt tttgtagttg aatttctctg 120
gctgatcttt tctgtacaga ttcatatatc tgcagagacg atatcattga ttatttgagc 180
ttcttttgaa ctatttcgtg taatttggga tgagagctct atgtatgtgt gtaaactttg 240
aagacaacaa gaaaggtaac aagtgaggga gggatgactc catgtcaaaa tagatgtcat 300
aagaggccca tcaataagtg cttgagccca ttagctagcc cagtaactac cagattgtga 360
gatggatgtg tgaacagttt ttttttgat gtaggactga aatgtgaaca acaggcgcat 420
gaaaggctaa attaggacaa tgataagcag aaataactta tcctctctaa cacttggcct 480
cacattgccc ttcacacaat ccacacacat ccaatcacaa cctcatcata tatctcccgc 540
taatcttttt ttctttgatc ttttttttt tgcttattat ttttttgact ttgatctccc 600
atcagttcat cttcttcttc ttcttctgat caacc 635

<210> 3
<211> 2038
<212> DNA
<213> *Arabidopsis thaliana*

<220>
<221> promoter
<222> (1)..(2038)
<223> TPT promoter

<400> 3
ataaccaaaa gaactctgat catgtacgta cccatttgcg tattccgccg ttgcggaatc 60
aaaaactgcc atagacttct ccacatcttt ttcagctcgc atcagtttga taacctgtga 120
aggcgtgatg ttctttgacc acttgaacat catcactttg ttacccatta ctaaacaatg 180
ataccttcca acaatagcaa agaatacacc tttttataga gaagaatctc agctacgcac 240
tacgtcgaac aggttgtgtg cataaacgat ttcgataggc ccaaaccaaa tgaaagaaac 300
acagaccaga aaaatcattt gatcttcaaa acatacgagt tccaaaagtg aaggaagcaa 360
caatgaaact cgttacgaac tagaaggtta atcaaattgc cggagaagaa tcgctcacca 420
gttttggcta gggtttatga aatgggagac tttagctgca aagagaagag tctctggacg 480
atttagaggg tgtctctcta ataggcaaca aagtacatat tattacagta ttaaccaaat 540
ttaaacgaat taagtgtcaa caaaagctta tataaaaaat ttaaagttta aaaattataa 600
aatatgtcaa caatatttta gtacttaaaa ttattatgcg aaatatttag atcaatggac 660
tactcatcta atatatttgc acctaatttt aaagtataaa ttcaaccaat aattagaaaa 720
tgatagctta tactcaaatt caacaagtta tatataaatg tatagatact acaatatcat 780
taacaaaagt caccttaaat aaatacacat atcttttatg ttctctattg ttttgcgtac 840
gctaacacaa tttctcatat gcaaaaggat gaatgagtaa caaattacct cataagaaca 900
atcatctttg cttacatact aatacaataa tcactcaatc aaccaataac atcaatcaca 960
taggtttaca tacaataatc actcaatcaa cttcataaga agaatcatgt ttacttaatt 1020
catcaattat ccccaaaaac accactatta agtataaact ataacatatt tgtagtgatg 1080
ggtcaacatt tttatcatat ttaaactcgg gttccctcaa atcgagaaat agtgaacatg 1140
taatattaat tttaaatcgc aattacagaa attaattgaa tttggtcaaa tggacagaat 1200
tttatagatt gggtggaact agaaaaaaaa aaaaaaagag tatagggtga attgagtaca 1260
tgaaagtaca tggtaatcct agttaaacgc ataatacatg tgggttcatt tgtatttttt 1320
tgtaacttac gagtaaactg gctacaacaa aaaaaattag aagatttttt tgttttgtag 1380
aaaaccctaa ttttagttat agttgtataa ctttgataaa attataaaat tgtattacga 1440
aaaaagtaat aagatattca aaaaagccta gaataacgta tatgactatg agcatgaaac 1500
tgcaagtcaa atgctgacag acaaccataa acaaaagaaa ttaaatagag atacctttaa 1560
aataagtaaa atttcattta taaaaaatct actttcttgt gaatctgtca cgttcaataa 1620

tttgaagacc actcaacata caaggtaaat aatgaaaaat aaaatctacc aaaatttcaa 1680
tcattattat cttccaaaaa aacaaaatta tacagatgat gatggtgata tggaacttcg 1740
attggctaat attcactgtg tctctaaaaa ccatccactt atcaagataa gatggaccct 1800
acactcatcc aatctaaacc agtatctcaa gattcttatc taattacatc attctctacc 1860
gttagatgaa attgaccatt aaccctacca taactccata caccgcgaga tactggatta 1920
accaaatcga gatcatcgta gccgtccgat caacaagtac catctcttga aatactcgaa 1980
atcctcataa gtccgtccct ctttgctctc actatcaaaa ctctgaattt cgatttca 2038

<210> 4
<211> 699
<212> DNA
<213> *Arabidopsis thaliana*

<220>
<221> promoter
<222> (1)..(699)
<223> FDS promoter

<400> 4
aacactctta acaccaaatc aacagattta acagcgaaat aagcttaagc ccatacaatt 60
atccgatcca aacaaatata atcgaaaccg gcagaggaat aagcaagtga atcaaaaagt 120
atgggacgag gaagaagatg atacctgaat gagaaagtca ataaccttga cccgaatcgt 180
tttgaagaaa atggagaaaa tcggttgtat ggaataaaat cttcgaatga tgagatatat 240
gatctctttg gtgtcagtca catggcacac gctatcaatt tagaaaaacg cggtggttgg 300
tcaccagaat tactacttct cggtctgatt tggtcatatc cgtattaagt ccggttaata 360
ttttccataa ctggggtttg aacattcggt ttctttttt cagttagtcc gatttggagt 420
tttgagtatg gaaaaataat actgaattta tttgttcaaa ctgttttgga aaaaatattt 480
cccttaatta cgaatataat taaaatttta aaattcattt tattagatct tggttaattc 540
ggtttaatgc attaatgaat ttcggtttaa gtcggttttc ggttttatg tcccaccact 600
atctacaacc gatgatcaac cttatctccg tattcaccac aaacagtcat cactctcact 660
tgacacaaaa actcttttgt ctccgtctct ctgtctctc 699

<210> 5
<211> 552
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: codon
adapted sequence for
phosphinotricin-N-acetyltransferase

<220>
<221> CDS
<222> (1)..(549)
<223> phosphinotricin-N-acetyltransferase (PAT)

<400> 5
atg tct ccg gag agg aga cca gtt gag att agg cca gct aca gca gcc 48
Met Ser Pro Glu Arg Arg Pro Val Glu Ile Arg Pro Ala Thr Ala Ala
1 5 10 15

gat atg gcc gcg gtt tgt gac atc gtt aac cat tac att gag acg tct 96
Asp Met Ala Ala Val Cys Asp Ile Val Asn His Tyr Ile Glu Thr Ser
20 25 30

aca gtg aac ttt agg aca gag cca caa aca cca caa gag tgg att gat 144
Thr Val Asn Phe Arg Thr Glu Pro Gln Thr Pro Gln Glu Trp Ile Asp
35 40 45

gac cta gag agg ttg caa gat aga tac cct tgg ttg gtt gct gag gtt 192
Asp Leu Glu Arg Leu Gln Asp Arg Tyr Pro Trp Leu Val Ala Glu Val
50 55 60

gag ggt gtt gtg gct ggt att gct tac gct ggg ccc tgg aag gct agg 240
Glu Gly Val Val Ala Gly Ile Ala Tyr Ala Gly Pro Trp Lys Ala Arg
65 70 75 80

aac gct tac gat tgg aca gtt gag agt act gtt tac gtg tca cat agg 288
Asn Ala Tyr Asp Trp Thr Val Glu Ser Thr Val Tyr Val Ser His Arg
85 90 95

cat caa agg ttg ggc cta gga tct aca ttg tac aca cat ttg ctt aag 336
His Gln Arg Leu Gly Leu Gly Ser Thr Leu Tyr Thr His Leu Leu Lys
100 105 110

tct atg gag gcg caa ggt ttt aag tct gtg gtt gct gtt ata ggc ctt 384
Ser Met Glu Ala Gln Gly Phe Lys Ser Val Val Ala Val Ile Gly Leu
115 120 125

cca aac gat cca tct gtt agg ttg cat gag gct ttg gga tac aca gcg 432
Pro Asn Asp Pro Ser Val Arg Leu His Glu Ala Leu Gly Tyr Thr Ala
130 135 140

cgg ggt aca ttg cgc gcg gct gga tac aag cat ggt gga tgg cat gat 480
Arg Gly Thr Leu Arg Ala Ala Gly Tyr Lys His Gly Gly Trp His Asp
145 150 155 160

gtt ggt ttt tgg caa agg gat ttt gag ttg cca gct cct cca agg cca 528
Val Gly Phe Trp Gln Arg Asp Phe Glu Leu Pro Ala Pro Pro Arg Pro
165 170 175

gtt agg cca gtt acc cag atc tga 552
Val Arg Pro Val Thr Gln Ile
180

<210> 6
<211> 183
<212> PRT
<213> Artificial sequence
<223> Description of the artificial sequence: codon adapted sequence for phosphinotricin-N-acetyltransferase

<400> 6
Met Ser Pro Glu Arg Arg Pro Val Glu Ile Arg Pro Ala Thr Ala Ala
1 5 10 15

Asp Met Ala Ala Val Cys Asp Ile Val Asn His Tyr Ile Glu Thr Ser
20 25 30

Thr Val Asn Phe Arg Thr Glu Pro Gln Thr Pro Gln Glu Trp Ile Asp
35 40 45

Asp Leu Glu Arg Leu Gln Asp Arg Tyr Pro Trp Leu Val Ala Glu Val
50 55 60

Glu Gly Val Val Ala Gly Ile Ala Tyr Ala Gly Pro Trp Lys Ala Arg
65 70 75 80

Asn Ala Tyr Asp Trp Thr Val Glu Ser Thr Val Tyr Val Ser His Arg
85 90 95

His Gln Arg Leu Gly Leu Gly Ser Thr Leu Tyr Thr His Leu Leu Lys
100 105 110

Ser Met Glu Ala Gln Gly Phe Lys Ser Val Val Ala Val Ile Gly Leu
115 120 125

Pro Asn Asp Pro Ser Val Arg Leu His Glu Ala Leu Gly Tyr Thr Ala
130 135 140

Arg Gly Thr Leu Arg Ala Ala Gly Tyr Lys His Gly Gly Trp His Asp
145 150 155 160

Val Gly Phe Trp Gln Arg Asp Phe Glu Leu Pro Ala Pro Pro Arg Pro
165 170 175

Val Arg Pro Val Thr Gln Ile
180

<210> 7
<211> 2013
<212> DNA
<213> *Arabidopsis thaliana*

<220>
<221> CDS
<222> (1)..(2010)
<223> acetolactate synthase

<400> 7
atg gcg gcg gca aca aca aca aca aca tct tct tcg atc tcc ttc 48
Met Ala Ala Ala Thr Thr Thr Thr Thr Ser Ser Ser Ile Ser Phe
1 5 10 15

tcc acc aaa cca tct cct tcc tcc tcc aaa tca cca tta cca atc tcc 96
Ser Thr Lys Pro Ser Pro Ser Ser Ser Lys Ser Pro Leu Pro Ile Ser
20 25 30

aga ttc tcc ctc cca ttc tcc cta aac ccc aac aaa tca tcc tcc tcc 144
Arg Phe Ser Leu Pro Phe Ser Leu Asn Pro Asn Lys Ser Ser Ser Ser
35 40 45

tcc cgc cgc cgc ggt atc aaa tcc agc tct ccc tcc tcc atc tcc gcc 192
Ser Arg Arg Arg Gly Ile Lys Ser Ser Ser Pro Ser Ser Ile Ser Ala
50 55 60

gtg ctc aac aca acc acc aat gtc aca acc act ccc tct cca acc aaa 240
Val Leu Asn Thr Thr Thr Asn Val Thr Thr Thr Pro Ser Pro Thr Lys
65 70 75 80

cct acc aaa ccc gaa aca ttc atc tcc cga ttc gct cca gat caa ccc 288
Pro Thr Lys Pro Glu Thr Phe Ile Ser Arg Phe Ala Pro Asp Gln Pro
85 90 95

cgc aaa ggc gct gat atc ctc gtc gaa gct tta gaa cgt caa ggc gta 336
Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg Gln Gly Val
100 105 110

gaa acc gta ttc gct tac cct gga ggt gca tca atg gag att cac caa 384
Glu Thr Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu Ile His Gln
115 120 125

gcc tta acc cgc tct tcc tca atc cgt aac gtc ctt cct cgt cac gaa 432
Ala Leu Thr Arg Ser Ser Ser Ile Arg Asn Val Leu Pro Arg His Glu
130 135 140

caa gga ggt gta ttc gca gca gaa gga tac gct cga tcc tca ggt aaa 480
Gln Gly Gly Val Phe Ala Ala Glu Gly Tyr Ala Arg Ser Ser Gly Lys
145 150 155 160

cca ggt atc tgt ata gcc act tca ggt ccc gga gct aca aat ctc gtt 528
Pro Gly Ile Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu Val
165 170 175

agc gga tta gcc gat gcg ttg tta gat agt gtt cct ctt gta gca atc 576
Ser Gly Leu Ala Asp Ala Leu Leu Asp Ser Val Pro Leu Val Ala Ile
180 185 190

aca gga caa gtc cct cgt cgt atg att ggt aca gat gcg ttt caa gag 624
Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala Phe Gln Glu
195 200 205

act ccg att gtt gag gta acg cgt tcg att acg aag cat aac tat ctt 672
Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His Asn Tyr Leu
210 215 220

gtg atg gat gtt gaa gat atc cct agg att att gag gaa gct ttc ttt 720
Val Met Asp Val Glu Asp Ile Pro Arg Ile Ile Glu Glu Ala Phe Phe
225 230 235 240

tta gct act tct ggt aga cct gga cct gtt ttg gtt gat gtt cct aaa 768
Leu Ala Thr Ser Gly Arg Pro Gly Pro Val Leu Val Asp Val Pro Lys
245 250 255

gat att caa caa cag ctt gcg att cct aat tgg gaa cag gct atg aga 816
Asp Ile Gln Gln Gln Leu Ala Ile Pro Asn Trp Glu Gln Ala Met Arg
260 265 270

tta cct ggt tat atg tct agg atg cct aaa cct ccg gaa gat tct cat 864
Leu Pro Gly Tyr Met Ser Arg Met Pro Lys Pro Pro Glu Asp Ser His
275 280 285

ttg gag cag att gtt agg ttg att tct gag tct aag aag cct gtg ttg 912
Leu Glu Gln Ile Val Arg Leu Ile Ser Glu Ser Lys Lys Pro Val Leu
290 295 300

tat gtt ggt ggt ggt tgt ttg aat tct agc gat gaa ttg ggt agg ttt 960
Tyr Val Gly Gly Gly Cys Leu Asn Ser Ser Asp Glu Leu Gly Arg Phe
305 310 315 320

gtt gag ctt acg ggg atc cct gtt gcg agt acg ttg atg ggg ctg gga 1008
Val Glu Leu Thr Gly Ile Pro Val Ala Ser Thr Leu Met Gly Leu Gly
325 330 335

tct tat cct tgt gat gat gag ttg tcg tta cat atg ctt gga atg cat 1056
Ser Tyr Pro Cys Asp Asp Glu Leu Ser Leu His Met Leu Gly Met His
340 345 350

ggg act gtg tat gca aat tac gct gtg gag cat agt gat ttg ttg ttg 1104
Gly Thr Val Tyr Ala Asn Tyr Ala Val Glu His Ser Asp Leu Leu Leu
355 360 365

gcg ttt ggg gta agg ttt gat gat cgt gtc acg ggt aag ctt gag gct 1152
Ala Phe Gly Val Arg Phe Asp Asp Arg Val Thr Gly Lys Leu Glu Ala
370 375 380

ttt gct agt agg gct aag att gtt cat att gat att gac tcg gct gag 1200
Phe Ala Ser Arg Ala Lys Ile Val His Ile Asp Ile Asp Ser Ala Glu
385 390 395 400

att ggg aag aat aag act cct cat gtg tct gtg tgt ggt gat gtt aag 1248
Ile Gly Lys Asn Lys Thr Pro His Val Ser Val Cys Gly Asp Val Lys
405 410 415

ctg gct ttg caa ggg atg aat aag gtt ctt gag aac cga gcg gag gag 1296
Leu Ala Leu Gln Gly Met Asn Lys Val Leu Glu Asn Arg Ala Glu Glu
420 425 430

ctt aag ctt gat ttt gga gtt tgg agg aat gag ttg aac gta cag aaa 1344
Leu Lys Leu Asp Phe Gly Val Trp Arg Asn Glu Leu Asn Val Gln Lys
435 440 445

cag aag ttt ccg ttg agc ttt aag acg ttt ggg gaa gct att cct cca 1392
Gln Lys Phe Pro Leu Ser Phe Lys Thr Phe Gly Glu Ala Ile Pro Pro
450 455 460

cag tat gcg att aag gtc ctt gat gag ttg act gat gga aaa gcc ata 1440
Gln Tyr Ala Ile Lys Val Leu Asp Glu Leu Thr Asp Gly Lys Ala Ile
465 470 475 480

ata agt act ggt gtc ggg caa cat caa atg tgg gcg gcg cag ttc tac 1488
Ile Ser Thr Gly Val Gly Gln His Gln Met Trp Ala Ala Gln Phe Tyr
485 490 495

aat tac aag aaa cca agg cag tgg cta tca tca gga ggc ctt gga gct 1536
Asn Tyr Lys Lys Pro Arg Gln Trp Leu Ser Ser Gly Gly Leu Gly Ala
500 505 510

atg gga ttt gga ctt cct gct gcg att gga gcg tct gtt gct aac cct 1584
Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ser Val Ala Asn Pro
515 520 525

gat gcg ata gtt gtg gat att gac gga gat gga agc ttt ata atg aat 1632
Asp Ala Ile Val Val Asp Ile Asp Gly Asp Gly Ser Phe Ile Met Asn
530 535 540

gtg caa gag cta gcc act att cgt gta gag aat ctt cca gtg aag gta 1680
Val Gln Glu Leu Ala Thr Ile Arg Val Glu Asn Leu Pro Val Lys Val
545 550 555 560

ctt tta tta aac aac cag cat ctt ggc atg gtt atg caa tgg gaa gat 1728
Leu Leu Leu Asn Asn Gln His Leu Gly Met Val Met Gln Trp Glu Asp
565 570 575

cgg ttc tac aaa gct aac cga gct cac aca ttt ctc ggg gat ccg gct 1776
Arg Phe Tyr Lys Ala Asn Arg Ala His Thr Phe Leu Gly Asp Pro Ala
580 585 590

cag gag gac gag ata ttc ccg aac atg ttg ctg ttt gca gca gct tgc 1824
Gln Glu Asp Glu Ile Phe Pro Asn Met Leu Leu Phe Ala Ala Ala Cys
595 600 605

ggg att cca gcg gcg agg gtg aca aag aaa gca gat ctc cga gaa gct 1872
Gly Ile Pro Ala Ala Arg Val Thr Lys Lys Ala Asp Leu Arg Glu Ala
610 615 620

att cag aca atg ctg gat aca cca gga cct tac ctg ttg gat gtg att 1920
Ile Gln Thr Met Leu Asp Thr Pro Gly Pro Tyr Leu Leu Asp Val Ile
625 630 635 640

tgt ccg cac caa gaa cat gtg ttg ccg atg atc ccg aat ggt ggc act 1968
Cys Pro His Gln Glu His Val Leu Pro Met Ile Pro Asn Gly Gly Thr
645 650 655

ttc aac gat gtc ata acg gaa gga gat ggc cgg att aaa tac tga 2013
Phe Asn Asp Val Ile Thr Glu Gly Asp Gly Arg Ile Lys Tyr
660 665 670

<210> 8
<211> 670
<212> PRT
<213> *Arabidopsis thaliana*

<400> 8
Met Ala Ala Ala Thr Thr Thr Thr Thr Thr Ser Ser Ser Ile Ser Phe
1 5 10 15

Ser Thr Lys Pro Ser Pro Ser Ser Ser Lys Ser Pro Leu Pro Ile Ser
20 25 30

Arg Phe Ser Leu Pro Phe Ser Leu Asn Pro Asn Lys Ser Ser Ser Ser
35 40 45

Ser Arg Arg Arg Gly Ile Lys Ser Ser Ser Pro Ser Ser Ile Ser Ala
50 55 60

Val Leu Asn Thr Thr Thr Asn Val Thr Thr Thr Pro Ser Pro Thr Lys
65 70 75 80

Pro Thr Lys Pro Glu Thr Phe Ile Ser Arg Phe Ala Pro Asp Gln Pro
85 90 95

Arg Lys Gly Ala Asp Ile Leu Val Glu Ala Leu Glu Arg Gln Gly Val
100 105 110

Glu Thr Val Phe Ala Tyr Pro Gly Gly Ala Ser Met Glu Ile His Gln
115 120 125

Ala Leu Thr Arg Ser Ser Ser Ile Arg Asn Val Leu Pro Arg His Glu
130 135 140

Gln Gly Gly Val Phe Ala Ala Glu Gly Tyr Ala Arg Ser Ser Gly Lys
145 150 155 160

Pro Gly Ile Cys Ile Ala Thr Ser Gly Pro Gly Ala Thr Asn Leu Val
165 170 175

Ser Gly Leu Ala Asp Ala Leu Leu Asp Ser Val Pro Leu Val Ala Ile
180 185 190

Thr Gly Gln Val Pro Arg Arg Met Ile Gly Thr Asp Ala Phe Gln Glu
195 200 205

Thr Pro Ile Val Glu Val Thr Arg Ser Ile Thr Lys His Asn Tyr Leu
210 215 220

Val Met Asp Val Glu Asp Ile Pro Arg Ile Ile Glu Glu Ala Phe Phe
225 230 235 240

Leu Ala Thr Ser Gly Arg Pro Gly Pro Val Leu Val Asp Val Pro Lys
245 250 255

Asp Ile Gln Gln Gln Leu Ala Ile Pro Asn Trp Glu Gln Ala Met Arg
260 265 270

Leu Pro Gly Tyr Met Ser Arg Met Pro Lys Pro Pro Glu Asp Ser His
275 280 285

Leu Glu Gln Ile Val Arg Leu Ile Ser Glu Ser Lys Lys Pro Val Leu
290 295 300

Tyr Val Gly Gly Gly Cys Leu Asn Ser Ser Asp Glu Leu Gly Arg Phe
305 310 315 320

Val Glu Leu Thr Gly Ile Pro Val Ala Ser Thr Leu Met Gly Leu Gly
325 330 335

Ser Tyr Pro Cys Asp Asp Glu Leu Ser Leu His Met Leu Gly Met His
340 345 350

Gly Thr Val Tyr Ala Asn Tyr Ala Val Glu His Ser Asp Leu Leu Leu
355 360 365

Ala Phe Gly Val Arg Phe Asp Asp Arg Val Thr Gly Lys Leu Glu Ala
370 375 380

Phe Ala Ser Arg Ala Lys Ile Val His Ile Asp Ile Asp Ser Ala Glu
385 390 395 400

Ile Gly Lys Asn Lys Thr Pro His Val Ser Val Cys Gly Asp Val Lys
405 410 415

Leu Ala Leu Gln Gly Met Asn Lys Val Leu Glu Asn Arg Ala Glu Glu
420 425 430

Leu Lys Leu Asp Phe Gly Val Trp Arg Asn Glu Leu Asn Val Gln Lys
435 440 445

Gln Lys Phe Pro Leu Ser Phe Lys Thr Phe Gly Glu Ala Ile Pro Pro
450 455 460

Gln Tyr Ala Ile Lys Val Leu Asp Glu Leu Thr Asp Gly Lys Ala Ile
465 470 475 480

Ile Ser Thr Gly Val Gly Gln His Gln Met Trp Ala Ala Gln Phe Tyr
485 490 495

Asn Tyr Lys Lys Pro Arg Gln Trp Leu Ser Ser Gly Gly Leu Gly Ala
500 505 510

Met Gly Phe Gly Leu Pro Ala Ala Ile Gly Ala Ser Val Ala Asn Pro
515 520 525

Asp Ala Ile Val Val Asp Ile Asp Gly Asp Gly Ser Phe Ile Met Asn
530 535 540

Val Gln Glu Leu Ala Thr Ile Arg Val Glu Asn Leu Pro Val Lys Val
545 550 555 560

Leu Leu Leu Asn Asn Gln His Leu Gly Met Val Met Gln Trp Glu Asp
565 570 575

Arg Phe Tyr Lys Ala Asn Arg Ala His Thr Phe Leu Gly Asp Pro Ala
580 585 590

Gln Glu Asp Glu Ile Phe Pro Asn Met Leu Leu Phe Ala Ala Ala Cys
595 600 605

Gly Ile Pro Ala Ala Arg Val Thr Lys Lys Ala Asp Leu Arg Glu Ala
610 615 620

Ile Gln Thr Met Leu Asp Thr Pro Gly Pro Tyr Leu Leu Asp Val Ile
625 630 635 640

Cys Pro His Gln Glu His Val Leu Pro Met Ile Pro Asn Gly Gly Thr
645 650 655

Phe Asn Asp Val Ile Thr Glu Gly Asp Gly Arg Ile Lys Tyr
660 665 670

<210> 9
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 9
gtcgacgaat tcgagagaca gagagacgg 29

<210> 10
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 10
gtcgacggta ccgattcaag cttcactgc 29

<210> 11
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 11
gagaattcga ttcaagcttc actgc 25

<210> 12
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 12
ccatgggaga gacagagaga cg 22

<210> 13
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 13
acggatccga gagacagaga gacggagaca aaa 33

<210> 14
<211> 32
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 14
gcggatccaa cactcttaac accaaatcaa ca 32

<210> 15
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 15
gtcgacggat ccggttgatc agaagaagaa gaagaagatg aact 44

<210> 16
<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 16
gtcgactcta gattcattat ttcgattttg atttcgtgac c 41

<210> 17
<211> 50
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 17
aagtcgacgg atccataacc aaaagaactc tgatcatgta cgtacccatt 50

<210> 18
<211> 50

<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 18
agacgtcgac tctagatgaa atcgaaattc agagttttga tagtgagagc 50

<210> 19
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 19
ccaaaccatg gtaagtttgt ctaaagctta 30

<210> 20
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 20
cggatccttt tgtgtttcgt cttctctcac g 31

<210> 21
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 21
gtctagaggc aaaccaccga gtgtt 25

<210> 22
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 22
cggtacctgt ttccagaaaa ttttgattca g 31

<210> 23
<211> 7554
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: binary plant expression vector

<400> 23
ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttattctaat aaacgctctt ttctcttagg tttacccgcc aatatatcct gtcaaacact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttgcat gcctgcaggt cgactctaga ctagtggatc cgatatcgcc 240
cgggctcgag gtaccgagct cgaattcact ggccgtcgtt ttacaacgac tcagagcttg 300
acaggaggcc cgatctagta acatagatga caccgcgcgc gataatttat cctagtttgc 360
gcgctatatt ttgttttcta tcgcgtatta aatgtataat tgcgggactc taatcataaa 420
aacccatctc ataaataacg tcatgcatta catgttaatt attacatgct taacgtaatt 480
caacagaaat tatatgataa tcatcgcaag accggcaaca ggattcaatc ttaagaaact 540
ttattgccaa atgtttgaac gatcggggat catccgggtc tgtggcggga actccacgaa 600
aatatccgaa cgcagcaaga tctaagcttg ggtcccgctc agaagaactc gtcaagaagg 660
cgatagaagg cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg 720
tcagcccatt cgccgccaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 780
tagcggtccg ccacacccag ccggccacag tcgatgaatc cagaaaagcg gccatttttcc 840
accatgatat tcggcaagca ggcatcgcca tgggtcacga cgagatcctc gccgtcgggc 900
atgcgcgcct tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc 960
agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt 1020
ttcgcttggt ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca 1080
tcagccatga tggatacttt ctcggcagga gcaaggtgag atgacaggag atcctgcccc 1140
ggcacttcgc ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct 1200
gcgcaaggaa cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca 1260
ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc 1320
cggaacacgg cggcatcaga gcagccgatt gtctgttgtg cccagtcata gccgaatagc 1380
ctctccaccc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac 1440
gatccagatc cggtgcagat tatttggatt gagagtgaat atgagactct aattggatac 1500
cgaggggaat ttatggaacg tcagtggagc atttttgaca agaaatattt gctagctgat 1560
agtgacctta ggcgactttt gaacgcgcaa taatggtttc tgacgtatgt gcttagctca 1620
ttaaactcca gaaacccgcg gctgagtggc tccttcaacg ttgcggttct gtcagttcca 1680
aacgtaaaac ggcttgtccc gcgtcatcgg cgggggtcat aacgtgactc ccttaattct 1740
ccgctcatga tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga 1800
tcctgcttgg taataattgt cattagattg tttttatgca tagatgcact cgaaatcagc 1860
caattttaga caagtatcaa acggatgtta attcagtaca ttaaagacgt ccgcaatgtg 1920
ttattaagtt gtctaagcgt caatttgttt acaccacaat atatcctgcc accagccagc 1980
caacagctcc ccgaccggca gctcggcaca aaatcaccac gcgttaccac cacgccggcc 2040
ggccgcatgg tgttgaccgt gttcgccggc attgccgagt tcgagcgttc cctaatcatc 2100
gaccgcaccc ggagcgggcg cgaggccgcc aaggcccgag gcgtgaagtt tggcccccgc 2160
cctaccctca ccccggcaca gatcgcgcac gcccgcgagc tgatcgacca ggaaggccgc 2220
accgtgaaag aggcggctgc actgcttggc gtgcatcgct cgaccctgta ccgcgcactt 2280

gagcgcagcg aggaagtgac gcccaccgag gccaggcggc gcggtgcctt ccgtgaggac 2340
gcattgaccg aggccgacgc cctggcggcc gccgagaatg aacgccaaga ggaacaagca 2400
tgaaaccgca ccaggacggc caggacgaac cgtttttcat taccgaagag atcgaggcgg 2460
agatgatcgc ggccgggtac gtgttcgagc cgcccgcgca cgtctcaacc gtgcggctgc 2520
atgaaatcct ggccggtttg tctgatgcca agctggcggc ctggccggcc agcttggccg 2580
ctgaagaaac cgagcgccgc cgtctaaaaa ggtgatgtgt atttgagtaa aacagcttgc 2640
gtcatgcggt cgctgcgtat atgatgcgat gagtaaataa acaaatacgc aaggggaacg 2700
catgaaggtt atcgctgtac ttaaccagaa aggcgggtca ggcaagacga ccatcgcaac 2760
ccatctagcc cgcgccctgc aactcgccgg ggccgatgtt ctgttagtcg attccgatcc 2820
ccagggcagt gcccgcgatt gggcggccgt gcgggaagat caaccgctaa ccgttgtcgg 2880
catcgaccgc ccgacgattg accgcgacgt gaaggccatc ggccggcgcg acttcgtagt 2940
gatcgacgga gcgcccccagg cggcggactt ggctgtgtcc gcgatcaagg cagccgactt 3000
cgtgctgatt ccggtgcagc caagccctta cgacatatgg gccaccgccg acctggtgga 3060
gctggttaag cagcgcattg aggtcacgga tggaaggcta caagcggcct ttgtcgtgtc 3120
gcgggcgatc aaaggcacgc gcatcggcgg tgaggttgcc gaggcgctgg ccgggtacga 3180
gctgcccatt cttgagtccc gtatcacgca gcgcgtgagc tacccaggca ctgccgccgc 3240
cggcacaacc gttcttgaat cagaacccga gggcgacgct gcccgcgagg tccaggcgct 3300
ggccgctgaa attaaatcaa aactcatttg agttaatgag gtaaagagaa aatgagcaaa 3360
agcacaaaca cgctaagtgc cggccgtccg agcgcacgca gcagcaaggc tgcaacgttg 3420
gccagcctgg cagacacgcc agccatgaag cgggtcaact ttcagttgcc ggcggaggat 3480
cacaccaagc tgaagatgta cgcggtacgc caaggcaaga ccattaccga gctgctatct 3540
gaatacatcg cgcagctacc agagtaaatg agcaaatgaa taaatgagta gatgaatttt 3600
agcggctaaa ggaggcggca tggaaaatca agaacaacca ggcaccgacg ccgtggaatg 3660
ccccatgtgt ggaggaacgg gcggttggcc aggcgtaagc ggctgggttg tctgccggcc 3720
ctgcaatggc actggaaccc ccaagcccga ggaatcggcg tgagcggtcg caaaccatcc 3780
ggcccggtac aaatcggcgc ggcgctgggt gatgacctgg tggagaagtt gaaggccgcg 3840
caggccgccc agcggcaacg catcgaggca gaagcacgcc ccggtgaatc gtggcaagcg 3900
gccgctgatc gaatccgcaa agaatcccgg caaccgccgg cagccggtgc gccgtcgatt 3960
aggaagccgc ccaagggcga cgagcaacca gatttttteg ttccgatgct ctatgacgtg 4020
ggcacccgcg atagtcgcag catcatggac gtggccgttt tccgtctgtc gaagcgtgac 4080
cgacgagctg gcgaggtgat ccgctacgag cttccagacg ggcacgtaga ggtttccgca 4140
gggccggccg gcatggccag tgtgtgggat tacgacctgg tactgatggc ggtttcccat 4200
ctaaccgaat ccatgaaccg ataccgggaa gggaagggag acaagcccgg ccgcgtgttc 4260
cgtccacacg ttgcggacgt actcaagttc tgccggcgag ccgatggcgg aaagcagaaa 4320
gacgacctgg tagaaacctg cattcggtta aacaccacgc acgttgccat gcagcgtacg 4380
aagaaggcca agaacggccg cctggtgacg gtatccgagg gtgaagcctt gattagccgc 4440
tacaagatcg taaagagcga aaccgggcgg ccggagtaca tcgagatcga gctagctgat 4500
tggatgtacc gcgagatcac agaaggcaag aacccggacg tgctgacggt tcaccccgat 4560
actttttga tcgatcccgg catcggccgt tttctctacc gcctggcacg ccgcgccgca 4620
ggcaaggcag aagccagatg gttgttcaag acgatctacg aacgcagtgg cagcgccgga 4680
gagttcaaga agttctgttt caccgtgcgc aagctgatcg ggtcaaatga cctgccggag 4740
tacgatttga aggaggaggc ggggcaggct ggcccgatcc tagtcatgcg ctaccgcaac 4800
ctgatcgagg gcgaagcatc cgccggttcc taatgtacgg agcagatgct agggcaaatt 4860
gccctagcag gggaaaaagg tcgaaaaggt ctctttcctg tggatagcac gtacattggg 4920
aacccaaagc cgtacattgg gaaccggaac ccgtacattg ggaacccaaa gccgtacatt 4980
gggaaccggt cacacatgta agtgactgat ataaaagaga aaaaaggcga tttttccgcc 5040

taaaactctt taaaacttat taaaactctt aaaaccccgcc tggcctgtgc ataactgtct 5100
ggccagcgca cagccgaaga gctgcaaaaa gcgcctaccc ttcggtcgct gcgctcccta 5160
cgccccgccg cttcgcgtcg gcctatcgcg gccgctggcc gctcaaaaat ggctggccta 5220
cggccaggca atctaccagg gcgcggacaa gccgccgcgt cgccactcga ccgccggcgc 5280
ccacatcaag gcaccctgcc tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat 5340
gcagctcccg gagacggtca cagcttgtct gtaagcggat gccgggagca gacaagcccg 5400
tcagggcgcg tcagcgggtg ttggcgggtg tcggggcgca gccatgaccc agtcacgtag 5460
cgatagcgga gtgtatactg gcttaactat gcggcatcag agcagattgt actgagagtg 5520
caccatatgc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggcgc 5580
tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 5640
tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 5700
aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 5760
tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 5820
tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 5880
cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 5940
agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 6000
tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc cttatccggt 6060
aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 6120
ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 6180
cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt 6240
accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt 6300
ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 6360
ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 6420
gtcatgcatg atatatctcc caatttgtgt agggcttatt atgcacgctt aaaaataata 6480
aaagcagact tgacctgata gtttggctgt gagcaattat gtgcttagtg catctaacgc 6540
ttgagttaag ccgcgccgcg aagcggcgtc ggcttgaacg aatttctagc tagacattat 6600
ttgccgacta ccttggtgat ctcgcctttc acgtagtgga caaattcttc caactgatct 6660
gcgcgcgagg ccaagcgatc ttcttcttgt ccaagataag cctgtctagc ttcaagtatg 6720
acgggctgat actgggccgg caggcgctcc attgcccagt cggcagcgac atccttcggc 6780
gcgattttgc cggttactgc gctgtaccaa atgcgggaca acgtaagcac tacatttcgc 6840
tcatcgccag cccagtcggg cggcgagttc catagcgtta aggtttcatt tagcgcctca 6900
aatagatcct gttcaggaac cggatcaaag agttcctccg ccgctggacc taccaaggca 6960
acgctatgtt ctcttgcttt tgtcagcaag atagccagat caatgtcgat cgtggctggc 7020
tcgaagatac ctgcaagaat gtcattgcgc tgccattctc caaattgcag ttcgcgctta 7080
gctggataac gccacggaat gatgtcgtcg tgcacaacaa tggtgacttc tacagcgcgg 7140
agaatctcgc tctctccagg ggaagccgaa gtttccaaaa ggtcgttgat caaagctcgc 7200
cgcgttgttt catcaagcct tacggtcacc gtaaccagca aatcaatatc actgtgtggc 7260
ttcaggccgc catccactgc ggagccgtac aaatgtacgg ccagcaacgt cggttcgaga 7320
tggcgctcga tgacgccaac tacctctgat agttgagtcg atacttcggc gatcaccgct 7380
tccccccatga tgtttaactt tgttttaggg cgactgccct gctgcgtaac atcgttgctg 7440
ctccataaca tcaaacatcg acccacggcg taacgcgctt gctgcttgga tgcccgaggc 7500
atagactgta ccccaaaaaa acagtcataa caagccatga aaaccgccac tgcg 7554

<210> 24
<211> 8327
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: binary plant expression vector

<400> 24
ttccatggac atacaaatgg acgaacggat aaacctttcc acgccctttt aaatatccga 60
ttattctaat aaacgctctt ttctcttagg tttacccgcc aatatatcct gtcaaacact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttgcat gccgatcccc cctgcagata gactatacta tgttttagcc 240
tgcctgctgg ctagctacta tgttatgtta tgttgtaaaa taaacacctg ctaaggtata 300
tctatctata ttttagcatg gctttctcaa taaattgtct ttccttatcg tttactatct 360
tatacctaat aatgaaataa taatatcaca tatgaggaac ggggcaggtt taggcatata 420
tatacgagtg tagggcggag tgggggggta cggtcgactc tagactagtg gatccnnntc 480
aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt aaagcacgag 540
gaagcggtca gcccattcgc cgccaagctc ttcagcaata tcacgggtag ccaacgctat 600
gtcctgatag cggtccgcca cacccagccg gccacagtcg atgaatccag aaaagcggcc 660
attttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga gatcctcgcc 720
gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc 780
ttcgtccaga tcatcctgat cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat 840
gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat gcagccgccg 900
cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg acaggagatc 960
ctgccccggc acttcgccca atagcagcca gtcccttccc gcttcagtga caacgtcgag 1020
cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 1080
cagttcattc agggcaccgg acaggtcggt cttgacaaaa agaaccgggc gcccctgcgc 1140
tgacagccgg aacacggcgg catcagagca gccgattgtc tgttgtgccc agtcatagcc 1200
gaatagcctc tccacccaag cggccggaga acctgcgtgc aatccatctt gttcaatcat 1260
nnnagatccg atatcgcccg ggctcgaggt accgagctcg aattcactgg ccgtcgtttt 1320
acaacgactc agccagcttg acaggaggcc cgatctagta acatagatga caccgcgcgc 1380
gataatttat cctagtttgc gcgctatatt ttgttttcta tcgcgtatta aatgtataat 1440
tgcgggactc taatcataaa aacccatctc ataaataacg tcatgcatta catgttaatt 1500
attacatgct taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca 1560
ggattcaatc ttaagaaact ttattgccaa atgtttgaac gatcgggggat catccgggtc 1620
tgtggcggga actccacgaa aatatccgaa cgcagcaaga tcggtcgatc gactcagatc 1680
tgggtaactg gcctaactgg ccttggagga gctggcaact caaaatcccc ttgccaaaaa 1740
ccaacatcat gccatccacc atgcttgtat ccagccgcgc gcaatgtacc ccgcgctgtg 1800
tatcccaaag cctcatgcaa cctaacagat ggatcgtttg gaaggcctat aacagcaacc 1860
acagacttaa aaccttgcgc ctccatagac ttaagcaaat gtgtgtacaa tgtagatcct 1920
aggcccaacc tttgatgcct atgtgacacg taaacagtac tctcaactgt ccaatcgtaa 1980
gcgttcctag ccttccaggg cccagcgtaa gcaataccag ccacaacacc ctcaacctca 2040
gcaaccaacc aagggtatct atcttgcaac ctctctaggt catcaatcca ctcttgtggt 2100
gtttgtggct ctgtcctaaa gttcactgta gacgtctcaa tgtaatggtt aacgatgtca 2160
caaaccgcgg ccatatcggc tgctgtagct ggcctaatct caactggtct cctctccgga 2220
gacatgtcga gattatttgg attgagagtg aatatgagac tctaattgga taccgagggg 2280
aatttatgga acgtcagtgg agcatttttg acaagaaata tttgctagct gatagtgacc 2340
ttaggcgact tttgaacgcg caataatggt ttctgacgta tgtgcttagc tcattaaact 2400
ccagaaaccc gcggctgagt ggctccttca acgttgcggt tctgtcagtt ccaaacgtaa 2460
aacggcttgt cccgcgtcat cggcgggggt cataacgtga ctcccttaat tctccgctca 2520

tgatcagatt gtcgtttccc gccttcagtt taaactatca gtgtttgaca ggatcctgct 2580
tggtaataat tgtcattaga ttgtttttat gcatagatgc actcgaaatc agccaatttt 2640
agacaagtat caaacggatg ttaattcagt acattaaaga cgtccgcaat gtgttattaa 2700
gttgtctaag cgtcaatttg tttacaccac aatatatcct gccaccagcc agccaacagc 2760
tccccgaccg gcagctcggc acaaaatcac cacgcgttac caccacgccg gccggccgca 2820
tggtgttgac cgtgttcgcc ggcattgccg agttcgagcg ttccctaatc atcgaccgca 2880
cccggagcgg gcgcgaggcc gccaaggccc gaggcgtgaa gtttggcccc cgccctaccc 2940
tcaccccggc acagatcgcg cacgcccgcg agctgatcga ccaggaaggc cgcaccgtga 3000
aagaggcggc tgcactgctt ggcgtgcatc gctcgaccct gtaccgcgca cttgagcgca 3060
gcgaggaagt gacgcccacc gaggccaggc ggcgcggtgc cttccgtgag gacgcattga 3120
ccgaggccga cgccctggcg gccgccgaga atgaacgcca agaggaacaa gcatgaaacc 3180
gcaccaggac ggccaggacg aaccgttttt cattaccgaa gagatcgagg cggagatgat 3240
cgcggccggg tacgtgttcg agccgcccgc gcacgtctca accgtgcggc tgcatgaaat 3300
cctggccggt ttgtctgatg ccaagctggc ggcctggccg gccagcttgg ccgctgaaga 3360
aaccgagcgc cgccgtctaa aaaggtgatg tgtatttgag taaaacagct tgcgtcatgc 3420
ggtcgctgcg tatatgatgc gatgagtaaa taaacaaata cgcaagggga acgcatgaag 3480
gttatcgctg tacttaacca gaaaggcggg tcaggcaaga cgaccatcgc aacccatcta 3540
gcccgcgccc tgcaactcgc cggggccgat gttctgttag tcgattccga tccccagggc 3600
agtgcccgcg attgggcggc cgtgcgggaa gatcaaccgc taaccgttgt cggcatcgac 3660
cgcccgacga ttgaccgcga cgtgaaggcc atcggccggc gcgacttcgt agtgatcgac 3720
ggagcgcccc aggcggcgga cttggctgtg tccgcgatca aggcagccga cttcgtgctg 3780
attccggtgc agccaagccc ttacgacata tgggccaccg ccgacctggt ggagctggtt 3840
aagcagcgca ttgaggtcac ggatggaagg ctacaagcgg cctttgtcgt gtcgcgggcg 3900
atcaaaggca cgcgcatcgg cggtgaggtt gccgaggcgc tggccgggta cgagctgccc 3960
attcttgagt cccgtatcac gcagcgcgtg agctacccag gcactgccgc cgccggcaca 4020
accgttcttg aatcagaacc cgagggcgac gctgcccgcg aggtccaggc gctggccgct 4080
gaaattaaat caaaactcat ttgagttaat gaggtaaaga gaaaatgagc aaaagcacaa 4140
acacgctaag tgccggccgt ccgagcgcac gcagcagcaa ggctgcaacg ttggccagcc 4200
tggcagacac gccagccatg aagcgggtca actttcagtt gccggcggag gatcacacca 4260
agctgaagat gtacgcggta cgccaaggca agaccattac cgagctgcta tctgaataca 4320
tcgcgcagct accagagtaa atgagcaaat gaataaatga gtagatgaat tttagcggct 4380
aaaggaggcg gcatggaaaa tcaagaacaa ccaggcaccg acgccgtgga atgccccatg 4440
tgtggaggaa cgggcggttg gccaggcgta agcggctggg ttgtctgccg gccctgcaat 4500
ggcactggaa ccccaagcc cgaggaatcg gcgtgagcgg tcgcaaacca tccggcccgg 4560
tacaaatcgg cgcggcgctg ggtgatgacc tggtggagaa gttgaaggcc gcgcaggccg 4620
cccagcggca acgcatcgag gcagaagcac gccccggtga atcgtggcaa gcggccgctg 4680
atcgaatccg caaagaatcc cggcaaccgc cggcagccgg tgcgccgtcg attaggaagc 4740
cgcccaaggg cgacgagcaa ccagattttt tcgttccgat gctctatgac gtgggcaccc 4800
gcgatagtcg cagcatcatg gacgtggccg ttttccgtct gtcgaagcgt gaccgacgag 4860
ctggcgaggt gatccgctac gagcttccag acgggcacgt agaggtttcc gcagggccgg 4920
ccggcatggc cagtgtgtgg gattacgacc tggtactgat ggcggtttcc catctaaccg 4980
aatccatgaa ccgataccgg gaagggaagg gagacaagcc cggccgcgtg ttccgtccac 5040
acgttgcgga cgtactcaag ttctgccggc gagccgatgg cggaaagcag aaagacgacc 5100
tggtagaaac ctgcattcgg ttaaacacca cgcacgttgc catgcagcgt acgaagaagg 5160
ccaagaacgg ccgcctggtg acggtatccg agggtgaagc cttgattagc cgctacaaga 5220
tcgtaaagag cgaaaccggg cggccggagt acatcgagat cgagctagct gattggatgt 5280

accgcgagat cacagaaggc aagaacccgg acgtgctgac ggttcacccc gattactttt 5340
tgatcgatcc cggcatcggc cgttttctct accgcctggc acgccgcgcc gcaggcaagg 5400
cagaagccag atggttgttc aagacgatct acgaacgcag tggcagcgcc ggagagttca 5460
agaagttctg tttcaccgtg cgcaagctga tcgggtcaaa tgacctgccg gagtacgatt 5520
tgaaggagga ggcggggcag gctggcccga tcctagtcat gcgctaccgc aacctgatcg 5580
agggcgaagc atccgccggt tcctaatgta cggagcagat gctagggcaa attgccctag 5640
caggggaaaa aggtcgaaaa ggtctctttc ctgtggatag cacgtacatt gggaacccaa 5700
agccgtacat tgggaaccgg aacccgtaca ttgggaaccc aaagccgtac attgggaacc 5760
ggtcacacat gtaagtgact gatataaaag agaaaaaagg cgatttttcc gcctaaaact 5820
ctttaaaact tattaaaact cttaaaaccc gcctggcctg tgcataactg tctggccagc 5880
gcacagccga agagctgcaa aaagcgccta cccttcggtc gctgcgctcc ctacgccccg 5940
ccgcttcgcg tcggcctatc gcggccgctg gccgctcaaa aatggctggc ctacggccag 6000
gcaatctacc agggcgcgga caagccgcgc cgtcgccact cgaccgccgg cgcccacatc 6060
aaggcaccct gcctcgcgcg tttcggtgat gacggtgaaa acctctgaca catgcagctc 6120
ccggagacgg tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc 6180
gcgtcagcgg gtgttggcgg gtgtcggggc gcagccatga cccagtcacg tagcgatagc 6240
ggagtgtata ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata 6300
tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgctcttccg 6360
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 6420
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 6480
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 6540
ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 6600
acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 6660
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 6720
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 6780
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 6840
gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 6900
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 6960
acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 7020
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 7080
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 7140
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatgc 7200
atgatatatc tcccaatttg tgtagggctt attatgcacg cttaaaaata ataaaagcag 7260
acttgacctg atagtttggc tgtgagcaat tatgtgctta gtgcatctaa cgcttgagtt 7320
aagccgcgcc gcgaagcggc gtcggcttga acgaatttct agctagacat tatttgccga 7380
ctaccttggt gatctcgcct ttcacgtagt ggacaaattc ttccaactga tctgcgcgcg 7440
aggccaagcg atcttcttct tgtccaagat aagcctgtct agcttcaagt atgacgggct 7500
gatactgggc cggcaggcgc tccattgccc agtcggcagc gacatccttc ggcgcgattt 7560
tgccggttac tgcgctgtac caaatgcggg acaacgtaag cactacattt cgctcatcgc 7620
cagcccagtc gggcggcgag ttccatagcg ttaaggtttc atttagcgcc tcaaatagat 7680
cctgttcagg aaccggatca aagagttcct ccgccgctgg acctaccaag gcaacgctat 7740
gttctcttgc ttttgtcagc aagatagcca gatcaatgtc gatcgtggct ggctcgaaga 7800
tacctgcaag aatgtcattg cgctgccatt ctccaaattg cagttcgcgc ttagctggat 7860
aacgccacgg aatgatgtcg tcgtgcacaa caatggtgac ttctacagcg cggagaatct 7920
cgctctctcc aggggaagcc gaagtttcca aaaggtcgtt gatcaaagct cgccgcgttg 7980
tttcatcaag ccttacggtc accgtaacca gcaaatcaat atcactgtgt ggcttcaggc 8040

cgccatccac tgcggagccg tacaaatgta cggccagcaa cgtcggttcg agatggcgct 8100
cgatgacgcc aactacctct gatagttgag tcgatacttc ggcgatcacc gcttcccccа 8160
tgatgtttaa ctttgtttta gggcgactgc cctgctgcgt aacatcgttg ctgctccata 8220
acatcaaaca tcgacccacg gcgtaacgcg cttgctgctt ggatgcccga ggcatagact 8280
gtaccccaaa aaaacagtca taacaagcca tgaaaaccgc cactgcg 8327

<210> 25
<211> 8441
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: binary
plant expression vector

<400> 25
ttccatggac atacaaatgg acgaacggat aaacctttтc acgccctttt aaatatccga 60
ttattctaat aaacgctctt ttctcttagg tttacccgcc aatatatcct gtcaaacact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaaaca gctatgacca 180
tgattacgcc aagcttgcat gccagcttga caggaggccc gatctagtaa catagatgac 240
accgcgcgcg ataatttatc ctagtttgcg cgctatattt tgttttctat cgcgtattaa 300
atgtataatt gcgggactct aatcataaaa acccatctca taaataacgt catgcattac 360
atgttaatta ttacatgctt aacgtaattc aacagaaatt atatgataat catcgcaaga 420
ccggcaacag gattcaatct taagaaactt tattgccaaa tgtttgaacg atcggggatc 480
atccgggtct gtggcgggaa ctccacgaaa atatccgaac gcagcaagat cggtcgatcg 540
actcagatct gggtaactgg cctaactggc cttggaggag ctggcaactc aaaatccctt 600
tgccaaaaac caacatcatg ccatccacca tgcttgtatc cagccgcgcg caatgtaccc 660
cgcgctgtgt atcccaaagc ctcatgcaac ctaacagatg gatcgtttgg aaggcctata 720
acagcaacca cagacttaaa accttgcgcc tccatagact taagcaaatg tgtgtacaat 780
gtagatccta ggcccaacct ttgatgccta tgtgacacgt aaacagtact ctcaactgtc 840
caatcgtaag cgttcctagc cttccagggc ccagcgtaag caataccagc cacaacaccc 900
tcaacctcag caaccaacca agggtatcta tcttgcaacc tctctaggtc atcaatccac 960
tcttgtggtg tttgtggctc tgtcctaaag ttcactgtag acgtctcaat gtaatggtta 1020
acgatgtcac aaaccgcggc catatcggct gctgtagctg gcctaatctc aactggtctc 1080
ctctccggag acatgtcgac tctagactag tggatccgat atcgcccggg ctcgaggtac 1140
cgagctcgaa ttcactggcc gtcgttttac aacgactcag agcttgacag gaggcccgat 1200
ctagtaacat agatgacacc gcgcgcgata atttatccta gtttgcgcgc tatattttgt 1260
tttctatcgc gtattaaatg tataattgcg ggactctaat cataaaaacc catctcataa 1320
ataacgtcat gcattacatg ttaattatta catgcttaac gtaattcaac agaaattata 1380
tgataatcat cgcaagaccg gcaacaggat tcaatcttaa gaaactttat tgccaaatgt 1440
ttgaacgatc ggggatcatc cgggtctgtg gcgggaactc cacgaaaata tccgaacgca 1500
gcaagatcta gagcttgggt cccgctcaga agaactcgtc aagaaggcga tagaaggcga 1560
tgcgctgcga atcgggagcg gcgataccgt aaagcacgag gaagcggtca gcccattcgc 1620
cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggtccgcca 1680
cacccagccg gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg 1740
gcaagcaggc atcgccatgg gtcacgacga gatcctcgcc gtcgggcatg cgcgccttga 1800
gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcatcctgat 1860
cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggt 1920
cgaatgggca ggtagccgga tcaagcgtat gcagccgccg cattgcatca gccatgatgg 1980

atactttctc ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgccca 2040
atagcagcca gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc 2100
ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg 2160
acaggtcggt cttgacaaaa agaaccgggc gcccctgcgc tgacagccgg aacacggcgg 2220
catcagagca gccgattgtc tgttgtgccc agtcatagcc gaatagcctc tccacccaag 2280
cggccggaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg 2340
tgcagattat ttggattgag agtgaatatg agactctaat tggataccga ggggaattta 2400
tggaacgtca gtggagcatt tttgacaaga aatatttgct agctgatagt gaccttaggc 2460
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520
acccgcggct gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc 2580
ttgtcccgcg tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca 2640
gattgtcgtt tcccgccttc agtttaaact atcagtgttt gacaggatcc tgcttggtaa 2700
taattgtcat tagattgttt ttatgcatag atgcactcga aatcagccaa ttttagacaa 2760
gtatcaaacg gatgttaatt cagtacatta aagacgtccg caatgtgtta ttaagttgtc 2820
taagcgtcaa tttgtttaca ccacaatata tcctgccacc agccagccaa cagctccccg 2880
accggcagct cggcacaaaa tcaccacgcg ttaccaccac gccggccggc cgcatggtgt 2940
tgaccgtgtt cgccggcatt gccgagttcg agcgttccct aatcatcgac cgcacccgga 3000
gcgggcgcga ggccgccaag gcccgaggcg tgaagtttgg ccccgccct accctcaccc 3060
cggcacagat cgcgcacgcc cgcgagctga tcgaccagga aggccgcacc gtgaaagagg 3120
cggctgcact gcttggcgtg catcgctcga ccctgtaccg cgcacttgag cgcagcgagg 3180
aagtgacgcc caccgaggcc aggcggcgcg gtgccttccg tgaggacgca ttgaccgagg 3240
ccgacgccct ggcggccgcc gagaatgaac gccaagagga acaagcatga aaccgcacca 3300
ggacggccag gacgaaccgt ttttcattac cgaagagatc gaggcggaga tgatcgcggc 3360
cgggtacgtg ttcgagccgc ccgcgcacgt ctcaaccgtg cggctgcatg aaatcctggc 3420
cggtttgtct gatgccaagc tggcggcctg gccggccagc ttggccgctg aagaaaccga 3480
gcgccgccgt ctaaaaaggt gatgtgtatt tgagtaaaac agcttgcgtc atgcggtcgc 3540
tgcgtatatg atgcgatgag taaataaaca aatacgcaag gggaacgcat gaaggttatc 3600
gctgtactta accagaaagg cgggtcaggc aagacgacca tcgcaaccca tctagcccgc 3660
gccctgcaac tcgccggggc cgatgttctg ttagtcgatt ccgatccccа gggcagtgcc 3720
cgcgattggg cggccgtgcg ggaagatcaa ccgctaaccg ttgtcggcat cgaccgcccg 3780
acgattgacc gcgacgtgaa ggccatcggc cggcgcgact tcgtagtgat cgacggagcg 3840
ccccaggcgg cggacttggc tgtgtccgcg atcaaggcag ccgacttcgt gctgattccg 3900
gtgcagccaa gcccttacga catatgggcc accgccgacc tggtggagct ggttaagcag 3960
cgcattgagg tcacggatgg aaggctacaa gcggcctttg tcgtgtcgcg ggcgatcaaa 4020
ggcacgcgca tcggcggtga ggttgccgag gcgctggccg ggtacgagct gcccattctt 4080
gagtcccgta tcacgcagcg cgtgagctac ccaggcactg ccgccgccgg cacaaccgtt 4140
cttgaatcag aacccgaggg cgacgctgcc cgcgaggtcc aggcgctggc cgctgaaatt 4200
aaatcaaaac tcatttgagt taatgaggta aagagaaaat gagcaaaagc acaaacacgc 4260
taagtgccgg ccgtccgagc gcacgcagca gcaaggctgc aacgttggcc agcctggcag 4320
acacgccagc catgaagcgg gtcaactttc agttgccggc ggaggatcac accaagctga 4380
agatgtacgc ggtacgccaa ggcaagacca ttaccgagct gctatctgaa tacatcgcgc 4440
agctaccaga gtaaatgagc aaatgaataa atgagtagat gaattttagc ggctaaagga 4500
ggcggcatgg aaaatcaaga acaaccaggc accgacgccg tggaatgccc catgtgtgga 4560
ggaacgggcg gttggccagg cgtaagcggc tgggttgtct gccggccctg caatggcact 4620
ggaaccccca agcccgagga atcggcgtga gcggtcgcaa accatccggc ccggtacaaa 4680
tcggcgcggc gctgggtgat gacctggtgg agaagttgaa ggccgcgcag gccgcccagc 4740

ggcaacgcat cgaggcagaa gcacgccccg gtgaatcgtg gcaagcggcc gctgatcgaa 4800
tccgcaaaga atcccggcaa ccgccggcag ccggtgcgcc gtcgattagg aagccgccca 4860
agggcgacga gcaaccagat tttttcgttc cgatgctcta tgacgtgggc acccgcgata 4920
gtcgcagcat catggacgtg gccgttttcc gtctgtcgaa gcgtgaccga cgagctggcg 4980
aggtgatccg ctacgagctt ccagacgggc acgtagaggt ttccgcaggg ccggccggca 5040
tggccagtgt gtgggattac gacctggtac tgatggcggt ttcccatcta accgaatcca 5100
tgaaccgata ccgggaaggg aagggagaca agcccggccg cgtgttccgt ccacacgttg 5160
cggacgtact caagttctgc cggcgagccg atggcggaaa gcagaaagac gacctggtag 5220
aaacctgcat tcggttaaac accacgcacg ttgccatgca gcgtacgaag aaggccaaga 5280
acggccgcct ggtgacggta tccgagggtg aagccttgat tagccgctac aagatcgtaa 5340
agagcgaaac cgggcggccg gagtacatcg agatcgagct agctgattgg atgtaccgcg 5400
agatcacaga aggcaagaac ccggacgtgc tgacggttca ccccgattac tttttgatcg 5460
atcccggcat cggccgtttt ctctaccgcc tggcacgccg cgccgcaggc aaggcagaag 5520
ccagatggtt gttcaagacg atctacgaac gcagtggcag cgccggagag ttcaagaagt 5580
tctgtttcac cgtgcgcaag ctgatcgggt caaatgacct gccggagtac gatttgaagg 5640
aggaggcggg gcaggctggc ccgatcctag tcatgcgcta ccgcaacctg atcgagggcg 5700
aagcatccgc cggttcctaa tgtacggagc agatgctagg gcaaattgcc ctagcagggg 5760
aaaaaggtcg aaaaggtctc tttcctgtgg atagcacgta cattgggaac ccaaagccgt 5820
acattgggaa ccggaacccg tacattggga acccaaagcc gtacattggg aaccggtcac 5880
acatgtaagt gactgatata aaagagaaaa aaggcgattt ttccgcctaa aactctttaa 5940
aacttattaa aactcttaaa acccgcctgg cctgtgcata actgtctggc cagcgcacag 6000
ccgaagagct gcaaaaagcg cctacccttc ggtcgctgcg ctccctacgc cccgccgctt 6060
cgcgtcggcc tatcgcggcc gctggccgct caaaaatggc tggcctacgg ccaggcaatc 6120
taccagggcg cggacaagcc gcgccgtcgc cactcgaccg ccggcgccca catcaaggca 6180
ccctgcctcg cgcgtttcgg tgatgacggt gaaaacctct gacacatgca gctcccggag 6240
acggtcacag cttgtctgta agcggatgcc gggagcagac aagcccgtca gggcgcgtca 6300
gcgggtgttg gcgggtgtcg gggcgcagcc atgacccagt cacgtagcga tagcggagtg 6360
tatactggct taactatgcg gcatcagagc agattgtact gagagtgcac catatgcggt 6420
gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgctct tccgcttcct 6480
cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa 6540
aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa 6600
aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc 6660
tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 6720
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc 6780
cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 6840
ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct 6900
gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 6960
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 7020
gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct 7080
acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa 7140
gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt tttttgttt 7200
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 7260
cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgcatgata 7320
tatctcccaa tttgtgtagg gcttattatg cacgcttaaa aataataaaa gcagacttga 7380
cctgatagtt tggctgtgag caattatgtg cttagtgcat ctaacgcttg agttaagccg 7440
cgccgcgaag cggcgtcggc ttgaacgaat ttctagctag acattatttg ccgactacct 7500

tggtgatctc gcctttcacg tagtggacaa attcttccaa ctgatctgcg cgcgaggcca 7560
agcgatcttc ttcttgtcca agataagcct gtctagcttc aagtatgacg ggctgatact 7620
gggccggcag gcgctccatt gcccagtcgg cagcgacatc cttcggcgcg attttgccgg 7680
ttactgcgct gtaccaaatg cgggacaacg taagcactac atttcgctca tcgccagccc 7740
agtcgggcgg cgagttccat agcgttaagg tttcatttag cgcctcaaat agatcctgtt 7800
caggaaccgg atcaaagagt tcctccgccg ctggacctac caaggcaacg ctatgttctc 7860
ttgcttttgt cagcaagata gccagatcaa tgtcgatcgt ggctggctcg aagatacctg 7920
caagaatgtc attgcgctgc cattctccaa attgcagttc gcgcttagct ggataacgcc 7980
acggaatgat gtcgtcgtgc acaacaatgg tgacttctac agcgcggaga atctcgctct 8040
ctccagggga agccgaagtt tccaaaaggt cgttgatcaa agctcgccgc gttgtttcat 8100
caagccttac ggtcaccgta accagcaaat caatatcact gtgtggcttc aggccgccat 8160
ccactgcgga gccgtacaaa tgtacggcca gcaacgtcgg ttcgagatgg cgctcgatga 8220
cgccaactac ctctgatagt tgagtcgata cttcggcgat caccgcttcc cccatgatgt 8280
ttaactttgt tttagggcga ctgccctgct gcgtaacatc gttgctgctc cataacatca 8340
aacatcgacc cacggcgtaa cgcgcttgct gcttggatgc ccgaggcata gactgtaccc 8400
caaaaaaaca gtcataacaa gccatgaaaa ccgccactgc g 8441

<210> 26
<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 26
aagtcgacgg atcctgatag cttatactca aattcaacaa gttat 45

<210> 27
<211> 1318
<212> DNA
<213> *Arabidopsis thaliana*

<220>
<221> promoter
<222> (1)..(1318)
<223> TPT truncated promoter

<400> 27
tgatagctta tactcaaatt caacaagtta tataaaatg tatagatact acaatatcat 60
taacaaaagt caccttaaat aaatacacat atcttttatg ttctctattg ttttgcgtac 120
gctaacacaa tttctcatat gcaaaaggat gaatgagtaa caaattacct cataagaaca 180
atcatctttg cttacatact aatacaataa tcactcaatc aaccaataac atcaatcaca 240
taggtttaca tacaataatc actcaatcaa cttcataaga agaatcatgt ttacttaatt 300
catcaattat ccccaaaaac accactatta agtataaact ataacatatt tgtagtgatg 360
ggtcaacatt tttatcatat ttaaactcgg gttcgctcaa atcgagaaat agtgaacatg 420
taatattaat tttaaatcgc aattacagaa attaattgaa tttggtcaaa tggacagaat 480
tttatagatt gggtggaact agaaaaaaaa aaaaaaagag tatagggtga attgagtaca 540
tgaaagtaca tggtaatcct agttaaacgc ataatacatg tgggttcatt tgtatttttt 600
tgtaacttac gagtaaactg gctacaacaa aaaaaattag aagatttttt tgttttgtag 660

aaaaccctaa ttttagttat agttgtataa ctttgataaa attataaaat tgtattacga 720
aaaaagtaat aagatattca aaaaagccta gaataacgta tatgactatg agcatgaaac 780
tgcaagtcaa atgctgacag acaaccataa acaaaagaaa ttaaatagag atacctttaa 840
aataagtaaa atttcattta taaaaaatct actttcttgt gaatctgtca cgttcaataa 900
tttgaagacc actcaacata caaggtaaat aatgaaaaat aaaatctacc aaaatttcaa 960
tcattattat cttccaaaaa aacaaaatta tacagatgat gatggtgata tggaacttcg 1020
attggctaat attcactgtg tctctaaaaa ccatccactt atcaagataa gatggaccct 1080
acactcatcc aatctaaacc agtatctcaa gattcttatc taattacatc attctctacc 1140
gttagatgaa attgaccatt aaccctacca taactccata caccgcgaga tactggatta 1200
accaaatcga gatcatcgta gccgtccgat caacaagtac catctcttga aatactcgaa 1260
atcctcataa gtccgtccct ctttgctctc actatcaaaa ctctgaattt cgatttca 1318

<210> 28
<211> 234
<212> DNA
<213> *Solanum tuberosum*

<220>
<221> terminator
<222> (1)..(234)
<223> terminator sequence of the Cathepsin D Inhibitor gene from potato

<400> 28
cctgcagata gactatacta tgttttagcc tgcctgctgg ctagctacta tgttatgtta 60
tgttgtaaaa taaacacctg ctaaggtata tctatctata ttttagcatg gctttctcaa 120
taaattgtct ttccttatcg tttactatct tatacctaat aatgaaataa taatatcaca 180
tatgaggaac ggggcaggtt taggcatata tatacgagtg tagggcggag tggg 234

<210> 29
<211> 298
<212> DNA
<213> *Vicia faba*

<220>
<221> terminator
<222> (1)..(298)
<223> terminator of storage protein gene VflElB3 from *Vicia faba*

<400> 29
gatcctgcaa tagaatgttg aggtgaccac tttctgtaat aaaataatta taaaataaat 60
ttagaattgc tgtagtcaag aacatcagtt ctaaaatatt aataaagtta tggccttttg 120
acatatgtgt ttcgataaaa aaatcaaaat aaattgagat ttattcgaaa tacaatgaaa 180
gtttgcagat atgagatatg tttctacaaa ataataactt aaaactcaac tatatgctaa 240
tgtttttctt ggtgtgtttc atagaaaatt gtatccgttt cttagaaaat gctcgtaa 298

We claim:

1. An expression cassette for transgenic expression of nucleic acids, comprising

a) a promoter according to SEQ ID NO: 1, 2 or 3 or

b) functional equivalents or equivalent fragments of a) which essentially possess the same promoter activities as a),

a) or b) being functionally linked to a nucleic acid sequence to be expressed transgenically.

2. An expression cassette as claimed in claim 1, wherein the functionally equivalent fragment is selected from the group of sequences described by SEQ ID NO: 4 and 27.

3. An expression cassette as claimed in either of claims 1 and 2, wherein

a) the nucleic acid sequence to be expressed is functionally linked to further genetic control sequences, or

b) the expression cassette contains additional functional elements, or

c) a) and b) apply.

4. An expression cassette as claimed in any of claims 1 to 3, wherein the nucleic acid sequence to be expressed transgenically enables

a) expression of a protein encoded by said nucleic acid sequence, or

b) expression of a sense or antisense RNA encoded by said nucleic acid sequence.

5. An expression cassette as claimed in any of claims 1 to 4, wherein the nucleic acid sequence to be expressed transgenically is selected from nucleic acids coding for selection markers, reporter genes, cellulases, chitinases, glucanases, ribosome-inactivating proteins, lysozymes, *Bacillus thuringiensis* endotoxin, α-amylase inhibitor, protease inhibitors, lectins, RNAases, ribozymes, acetyl-CoA

carboxylases, phytases, the 2S albumin from *Bertholletia excelsa*, antifreeze proteins, trehalose phosphate synthase, trehalose phosphate phosphatase, trehalase, DREB1A factor, farnesyl transferases, ferritin, oxalate oxidase, calcium-dependent protein kinases, calcineurins, glutamate dehydrogenases, the N-hydroxylating multifunctional cytochrome P-450, the transcriptional activator CBF1, phytoene desaturases, polygalacturonases, flavonoid 3'-hydroxylases, dihydroflavanol 4-reductases, chalcone isomerases, chalcone synthases, flavanone 3-beta-hydroxylases, flavone synthase II, branching enzyme Q, starch branching enzyme.

6. An expression cassette as claimed in any of claims 1 to 5, wherein the nucleic acid sequence to be expressed transgenically is selected from the group consisting of the nucleic acid sequences with the GenBank accession numbers U77378, AF306348, A19451, L25042, S78423, U32624, X78815, AJ002399, AF078796, AB044391, AJ222980, X14074, AB045593, AF017451, AF276302, AB061022, X72592, AB045592, AR123356.

7. An expression cassette as claimed in any of claims 1 to 4, wherein the nucleic acid sequence to be expressed transgenically is selected from the group consisting of positive selection markers, negative selection markers and factors which give a growth advantage.

8. An expression cassette as claimed in claim 7, wherein the selection marker is selected from the group consisting of proteins which impart a resistance to antibiotics, metabolism inhibitors, herbicides or biocides.

9. An expression cassette as claimed in either of claims 7 and 8, wherein the selection marker is selected from the group consisting of proteins, which impart a resistance to phosphinothricin, glyphosate, bromoxynil, dalapon, 2-deoxyglucose 6-phosphate, tetracyclines, ampicillin, kanamycin, G 418, neomycin, paromomycin, bleomycin, zeocin, hygromycin, chloramphenicol, sulfonyl urea herbicides, imidazolinone herbicides.

10. A transgenic expression cassette as claimed in any of claims 7 to 9, wherein the selection marker is selected from the group consisting of phosphinothricin acetyltransferases, 5-enolpyruvylshikimate 3-phosphate synthases, glyphosate oxidoreductases, dehalogenases, nitrilases, neomycin phosphotransferases, DOG^R1 genes, acetolactate synthases,

hygromycin phosphotransferases, chloramphenicol acetyltransferases, streptomycin adenylyltransferases, β-lactamases, tetA genes, tetR genes, isopentenyl transferases, thymidine kinases, diphtheria toxin A, cytosine deaminase (codA), cytochrome P450, haloalkane dehalogenases, iaaH gene, tms2 gene, β-glucuronidases, mannose 6-phosphate isomerases, UDP-galactose 4-epimerases.

11. A transgenic expression cassette as claimed in any of claims 7 to 10, wherein the selection marker is encoded by nucleic acid sequences

i) described by SEQ ID NO: 5 or 6, or

ii) described by or comprised in the sequences described by GenBank Acc.-No.: X17220, X05822, M22827, X65195, AJ028212, X17220, X05822 M22827, X65195, AJ028212, X63374, M10947, AX022822, AX022820, E01313, J03196, AF080390, AF234316, AF080389, AF234315, AF234314, U00004, NC001140, X51514, AB049823, AF094326, X07645, X07644, A19547, A19546, A19545, I05376, I05373, X74325, AF294981, AF234301, AF234300, AF234299, AF234298, AF354046, AF354045, X65876, X51366, AJ278607, L36849, AB025109, AL133315.

12. A vector comprising an expression cassette as claimed in any of claims 1 to 11.

13. A method for transgenic expression of nucleic acids, wherein a nucleic acid sequence which is functionally linked to

a) a promoter according to SEQ ID NO: 1, 2 or 3 or

b) a functional equivalent or equivalent fragment of a) which essentially possesses the same promoter activities as a),

is expressed transgenically.

14. A method as claimed in claim 13, wherein the functionally equivalent fragment is selected from the group of sequences described by SEQ ID NO: 4 and 27.

15. A method as claimed in any of claims 13 or 14, wherein

a) the nucleic acid sequence to be expressed is functionally linked to further genetic control sequences, or

b) an expression cassette used contains additional functional elements, or

c) a) and b) apply.

16. A method as claimed in any of claims 13 to 15, wherein the nucleic acid sequence to be expressed transgenically enables

a) expression of a protein encoded by said nucleic acid sequence, or

b) expression of a sense or antisense RNA encoded by said nucleic acid sequence.

17. A method as claimed in any of claims 13 to 16, wherein the nucleic acid sequence to be expressed transgenically is selected from nucleic acids coding for selection markers, reporter genes, cellulases, chitinases, glucanases, ribosome-inactivating proteins, lysozymes, *Bacillus thuringiensis* endotoxin, α-amylase inhibitor, protease inhibitors, lectins, RNAases, ribozymes, acetyl-CoA carboxylases, phytases, the 2S albumin from *Bertholletia excelsa*, antifreeze proteins, trehalose phosphate synthase, trehalose phosphate phosphatase, trehalase, DREB1A factor, farnesyl transferases, ferritin, oxalate oxidase, calcium-dependent protein kinases, calcineurins, glutamate dehydrogenases, the N-hydroxylating multifunctional cytochrome P-450, the transcriptional activator CBF1, phytoene desaturases, polygalacturonases, flavonoid 3'-hydroxylases, dihydroflavanol 4-reductases, chalcone isomerases, chalcone synthases, flavanone 3-beta-hydroxylases, flavone synthase II, branching enzyme Q, starch branching enzyme.

18. A method as claimed in any of claims 13 to 17, wherein the nucleic acid sequence to be expressed transgenically is selected from the group consisting of the nucleic acid sequences with the GenBank accession numbers U77378, AF306348, A19451, L25042, S78423, U32624, X78815, AJ002399, AF078796, AB044391, AJ222980, X14074, AB045593, AF017451, AF276302, AB061022, X72592, AB045592, AR123356.

19. A method for selecting transformed organisms, wherein a nucleic acid sequence coding for a selection marker, which is functionally and transgenically linked to

a) a promoter according to SEQ ID NO: 1, 2 or 3, or

b) a functional equivalent or equivalent fragment of a) which essentially possesses the same promoter activities as a),

is introduced into an organism, the selection marker is expressed and a selection is carried out.

20. A method as claimed in claim 19, wherein the selection marker is selected from the groups of selection markers which are mentioned in claims 6, 7, 8, 9 and 10.

21. A transgenic organism transformed with an expression cassette as claimed in claims 1 to 11 or with a vector as claimed in claim 12.

22. A transgenic organism as claimed in claim 21 selected from the group consisting of bacteria, yeasts, fungi, animal and plant organisms.

23. A transgenic organism as claimed in either of claims 21 and 22 selected from the group consisting of *Arabidopsis*, tomato, tobacco, potatoes, corn, oilseed rape, wheat, barley, sunflowers, millet, beet, rye, oats, sugarbeet, bean plants and soyabean.

24. A cell culture, plant or transgenic propagation material, derived from a transgenic organism as claimed in any of claims 21 to 23.

25. The use of a transgenic organism as claimed in any of claims 21 to 23 or of cell cultures, parts or transgenic propagation material derived therefrom as claimed in claim 24 for the production of food- and feedstuffs, seed, pharmaceuticals or fine chemicals.

26. The use as claimed in claim 25, wherein the fine chemicals are enzymes, vitamins, amino acids, sugars, saturated or unsaturated fatty acids, natural or synthetic flavorings, aromatizing substances or colorants.

27. The use as claimed in claim 25, wherein the pharmaceutical is an antibody, enzyme or pharmaceutically active protein.

28. A method for preparing pharmaceuticals or fine chemicals in transgenic organisms as claimed in any of claims 21 to 23 or in cell cultures, parts or transgenic propagation material derived therefrom as claimed in claim 24, which comprises growing the transgenic organism and isolating the desired pharmaceutical or the desired fine chemical.

29. A method as claimed in claim 28, wherein the fine chemicals are enzymes, vitamins, amino acids, sugars, saturated or unsaturated fatty acids, natural or synthetic flavorings, aromatizing substances or colorants.

30. A method as claimed in claim 28, wherein the pharmaceutical is an antibody, enzyme or pharmaceutically active protein.

Application number / numéro de demande: EP02/07527

Figures: 1A,B,C/2,A,B/5

Pages:

Unscannable items
received with this application
(Request original documents in File Prep. Section on the 10th floor)

Documents reçu avec cette demande ne pouvant être balayés
(Commander les documents originaux dans la section de préparation des dossiers au 10ème étage)

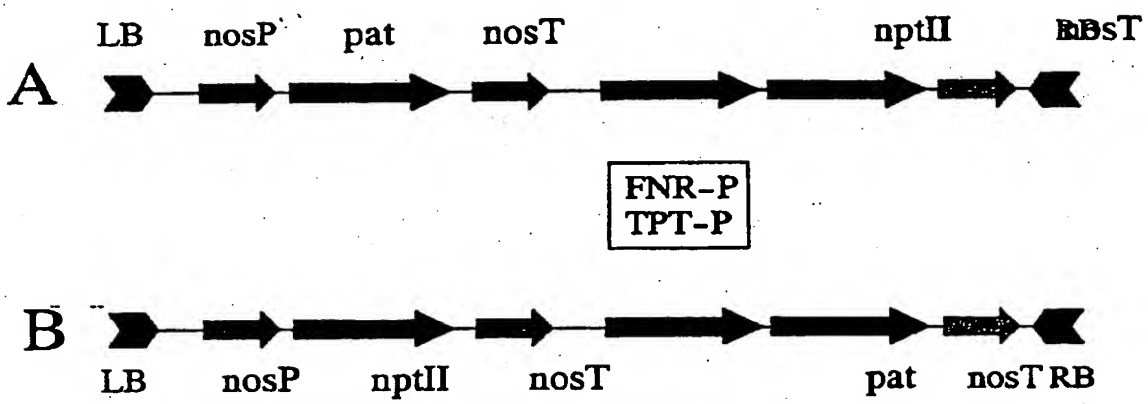

A
LB
nosP
pat
nosT
nptII
nosT
FNR-P
TPT-P
B
LB
nosP
nptII
nosT
pat
nosT
RB

Fig. 3